Racial Disparities in Federal Sentencing: Evidence from Drug Mandatory Minimums*

Cody Tuttle

February 19, 2023

Abstract

I study racial disparities in the criminal justice system by analyzing abnormal bunching in the distribution of crack-cocaine amounts used in federal sentencing. I compare cases sentenced before and after the Fair Sentencing Act, a 2010 law that changed the 10-year mandatory minimum threshold for crack-cocaine from 50g to 280g. First, I find that after 2010, there is a sharp increase in the fraction of cases sentenced at 280g (the point that now triggers a 10-year mandatory minimum), and that this increase is disproportionately large for black and Hispanic offenders. I then explore several possible explanations for the observed racial disparities, including racial discrimination that occurs after entry into the criminal justice system. I analyze data from multiple stages in the criminal justice system and find that the increased bunching for minority offenders is driven by prosecutorial discretion, specifically as used by about 20-30% of prosecutors. Moreover, the fraction of cases at 280g falls in 2013 when evidentiary standards become stricter. Finally, the racial disparity in the increase cannot be explained by differences in education, sex, age, criminal history, seized drug amount, or other elements of the crime, but it can be largely explained by a measure of state-level racial animus. These results shed light on the role of prosecutorial discretion and racial discrimination as causes of racial disparities in sentencing.

JEL Classification: J15, K14, K41, K42

Keywords: Racial disparities, discrimination, mandatory minimums, prosecutors, sentencing

*Affiliation: UT Austin, Department of Economics. Email: cody.tuttle@utexas.edu. I thank Melissa Kearney for her continued guidance and support. For helpful comments, I thank Alex Albright, David Arnold, Marika Cabral, Jonathan Caulkins, Prateik Dalmia, Stephanie Didwania, Allan Drazen, Mike Geruso, Jessica Goldberg, John Haltiwanger, Sara Heller, Judy Hellerstein, Brian Johnson, Ethan Kaplan, John Laub, Kyung Park, Nolan Pope, Stephanie Rennane, Matthew Staiger, Lesley Turner, Mateo Uribe-Castro, Sergio Urzua, Stan Veuger, Riley Wilson, Crystal Yang, and the participants at the University of Maryland Microeconomics Workshop, the University of Maryland Political Economy Workshop, the APPAM 2018 DC Regional Student Conference, the APPAM 2018 Fall Research Conference, the Population Association of America 2019 Conference, the Society of Labor Economists 2019 Conference, the Western Economics Association International Graduate Student Workshop, University of Chicago Crime Lab, DC Political Economy Workshop, the Southern Economic Association 2019 Conference, GMU Schar School of Policy and Government, Wellesley College, the FDIC, Yale School of Management, Yale University, Chicago Booth School of Business, Northwestern Kellogg School of Management, Chicago Harris School of Public Policy, Georgetown McCourt School of Public Policy, the Criminal Justice Administrative Records System (CJARS) at University of Michigan, Florida State University, Southern Methodist University, University of Texas at Austin, University of California San Diego, Claremont Graduate University, the ACLU, and the RAND Drug Policy Research Center. Finally, I am grateful to Robert Breetz and several other unnamed attorneys for feedback and insights about the institutional setting.

Racial differences in sentencing are a persistent concern in America. In recent federal cases, black offenders face sentences that are 20 percent longer than the sentences handed down for white offenders (USSC 2017). These added years are costly for society at large and for the people incarcerated. The Bureau of Prisons estimates the direct care cost of incarcerating a person is about $11,000 (in 2015 dollars) per year (US DOJ 2011). Mueller-Smith (2015) estimates an additional year in prison for felony defendants causes a 30 percent decrease in formal earnings post-release and significant lost wages while incarcerated. Even more, those incarcerated must confront serious physical and psychological costs of prison, in addition to the more intangible cost of their lost freedom (Haney 2001; The Hamilton Project 2016; BOP 2020). Due to racial sentencing disparities, these costs are disproportionately borne by black and Hispanic offenders.[1] For policy to confront these disparities, we must understand the root causes. One explanation for disparate sentences is that people of different races are different *upon entry* into the criminal justice system. Another explanation, however, is that *after entry* into the system, people are treated differently by race. Critically, these two explanations are not mutually exclusive, and any differences upon entry may themselves be a product of discrimination embedded in other US institutions. Still, understanding where and why disparities arise is crucial for informing the policy response.

In this paper, I examine racial sentencing disparities and test the second explanation: that agents in the criminal justice system (police, prosecutors, judges, etc.) treat black and Hispanic defendants differently than similar white defendants.[2] I focus on federal crack-cocaine cases and the application of mandatory minimum sentences in the US legal system to provide new evidence on this broader question. While concerns about disparities across groups and the discretion of criminal justice agents are underscored by the size of the criminal justice system in the US, these issues arise in legal systems across the globe (Council of Europe 1992; EMCDDA 2009; Tonry 2012; Nagrecha 2021). In many countries, prosecutors, in particular, have power over key decisions related to case dismissal, diversion, and charging (Tonry 2012; Mueller-Smith and Schnepel 2021). The specific setting in the US is also independently important. Approximately 20% of federal drug cases involve a crack-cocaine offense, and racial sentencing gaps are particularly large in these cases. In addition, the structure of mandatory minimums and recent changes in crack-cocaine mandatory minimums provide a unique opportunity to study discretion and racial disparities in the justice system.

In federal drug trafficking cases, a mandatory minimum sentence is triggered if the drug trafficking involves an amount of drugs equal to or above a threshold amount. Legal rules about police sting operations and the type of evidence admissible in federal court give both police and prosecutors power to influence the amount used in sentencing. If police or prosecutors want to increase the likelihood of a harsh sentence, they can use their discretion to move the amount of drugs to the threshold amount or just above it. Such quantity

[1] In the USSC variable **newrace**, four values are recorded for the offender's "race"–(1) non-Hispanic white, (2) non-Hispanic black, (3) Hispanic, and (4) other. Throughout the paper, I frequently use the term "race" in reference to Hispanic ethnicity to be consistent with this terminology used in the data and with this categorization that interacts racial and ethnic divisions.

[2] I use the term "offender" to describe a person in the final sentencing data or a person who has committed an offense (e.g. when talking about "offender responses" to the Fair Sentencing Act). Otherwise, I use the term "defendant."

thresholds are common in the US legal system and abroad. Drápal and Šoltés (2021) study how quantity thresholds influence decisions made by Czech prosecutors and detail the ways that many court systems across Europe employ thresholds in sentencing. I describe in detail how prosecutors in the US can influence the recorded drug amount in Section II.A.2. This paper studies whether police or prosecutors respond to this sentencing incentive and whether their responses are racially disparate. Specifically, I test for an excess mass or concentration (i.e. bunching) of cases at and above the mandatory minimum threshold (i.e. the use of discretion to increase the likelihood of a harsh sentence) and for differences in the concentration of cases by race (i.e. a racial disparity in the use of discretion).

With the Fair Sentencing Act in 2010, the 10-year mandatory minimum threshold for crack-cocaine was increased from 50g (i.e. 50 grams) to 280g.[3] This paper focuses on crack-cocaine because it is the only drug for which the federal mandatory minimum threshold has changed since the adoption of mandatory minimums in the 1980s. The shift to 280g is especially useful empirically since the new threshold is set at a point with zero bunching prior to 2010. All other mandatory minimum thresholds are set at somewhat natural bunching points (e.g. 50g) that do not vary over time.[4] Nevertheless, Table A2 shows that disparities in bunching exist for all major drugs and discusses the difficulties in interpreting bunching in other drugs.

Using this time variation in the mandatory minimum threshold, I implement a difference-in-bunching design where I first assume the pre-2010 distribution of drug amounts is a good counterfactual for what the post-2010 distribution would look like with the pre-2010 thresholds (Kleven 2016). I find the fraction of cases bunched at and above 280g increases after 2010, and that the increase is much larger for black and Hispanic offenders than for white offenders. To be clear, this is not intended as an evaluation of the Fair Sentencing Act, which likely caused a decline in sentences after 2010 (USSC 2015a; Bjerk 2017a). Rather, these results imply that police or prosecutors dampened the effect of the Fair Sentencing Act by increasing the drug amount charged for some defendants.[5] Also, these results do not imply that the use of discretion or the racial disparity in the use of discretion began after 2010. I take the shift to 280g as an opportunity to detect these behaviors that are otherwise difficult to detect.

After documenting this increase in bunching at 280g and the racial disparity in that increase, the paper proceeds with three results to put this key finding in context and provide evidence on channels. First, I rule out the possibility that the racial disparity in bunching at 280g is caused by racial differences in drug involvement. Second, I use data at multiple stages in the criminal justice process to estimate who

[3]The Fair Sentencing Act also shifted the 5-year threshold from 5g to 28g. I focus on the 10-year threshold for two reasons. First, 28g is below the pre-2010 10-year mandatory minimum threshold of 50g–this yields incentives for prosecutors to shift cases that would have been charged both above 50g and cases that would have been charged below 28g into the 28-50g range. Second, estimating whether the disparity in bunching is conditional on underlying drug involvement requires a range below the threshold that is not subject to strategic sentencing incentives. This is a reasonable assumption for the 60-280g range pre-2010, but would not be a reasonable assumption for the 6-28g range pre-2010 because those cases may be bunched at 50g.

[4]These amounts exhibit bunching even when they are not thresholds, likely a "round number" bias by police, prosecutors, etc.

[5]I use the term "charged" to mean the person was alleged to be involved with that amount. Prior to 2013, prosecutors did not have to charge drug amounts as part of the indictment; it was a factual determination relegated to the sentencing phase.

is responsible for the bunching of cases at 280g, and I find that it is a result of prosecutor decisions. I supplement my findings with tests related to the role of prosecutors, including an examination of a Supreme Court case that altered evidentiary standards specifically in mandatory minimum cases. Third, I explore whether the disparity can be attributed to discrimination from prosecutors–I rule out several alternative explanations and provide suggestive evidence of taste-based discrimination. I preview these results below.

The first result of similar drug involvement by race is rooted in evidence that cases bunched at 280g are primarily cases that would have been charged with 60-280g prior to 2010. I show that the pre-2010 distributions from 60-280g are nearly identical by race. Under the assumption that these distributions reflect true drug involvement, their similarity by race indicates similar drug involvement by race. If these pre-2010 amounts are biased against black and Hispanic offenders, then this would imply that black and Hispanic offenders have lower drug involvement than white offenders. The estimated disparity in bunching at 280g would then be an underestimate of the disparity conditional on drug involvement. This evidence on drug involvement is further supported by data on drug seizures which show similar seized amounts by race and data from inmate surveys which show similar reported involvement by race among federal inmates.

Next, using survey data on drug use and selling, data from state-level drug convictions, and drug seizure records, I rule out the possibilities that the bunching is a result of offender responses to the Fair Sentencing Act, a shifting of cases from state to federal court after 2010, or strategic responses from police or federal agents. Since I do not attribute the rise in bunching at 280g to one of these earlier stages, I turn my attention to prosecutors. Prosecutors can legally influence the drug quantity involved in an offense because the quantity of drugs used to determine sentencing is not strictly tied to the quantity found on the offender at the time of arrest (USSC 2015b). I find bunching at 280g after 2010 in case management data from the Executive Office of the US Attorney. I also find that approximately 30% of prosecutors are responsible for the rise in cases with 280g after 2010, and that there is variation in prosecutor-level bunching both within and between districts. These results, among others detailed in Section V.B.4, suggest that the observed bunching at sentencing is specifically due to prosecutorial discretion.

The US Supreme Court issued a 5-4 decision in *Alleyne v. United States* on June 17, 2013 that changed the evidentiary standard necessary for facts that raise a defendant's exposure to mandatory minimums (Bala 2015). Previously, prosecutors could present evidence on drug quantities to the presiding judge, and the judge would decide, based on the preponderance of evidence, whether the mandatory minimum applied. The Supreme Court ruling in *Alleyne* requires that prosecutors present this evidence to the jury, which evaluates it based on the stricter "beyond a reasonable doubt" standard.[6] The case management data from the Executive Office of the US Attorney show that from 2011-2013, 9.1% of cases were recorded in the 280-290g range. From 2014-2016, however, 6.8% of cases were recorded as 280-290g. Using a

[6]Although the vast majority of cases end in a plea, plea bargains happen in the "shadow" of the trial or the judge (LaCasse and Payne 1999, Bibas 2004). By changing the evidentiary standard at trial, *Alleyne* changes the prosecutor's bargaining position and affects even those cases that end in a plea.

difference-in-discontinuities design, I show that the practice of bunching ballooned in the run up to *Alleyne*, and that it was partially reined in by the Supreme Court decision. This suggests prosecutors were submitting evidence under the judicial fact-finding system that would not hold up under the scrutiny of a jury.

After documenting a racial disparity in bunching at 280g and studying the role of prosecutorial discretion in producing that disparity, I explore whether the disparity can be attributed to discrimination from prosecutors. I consider four potential sources of the racial disparity. First, I explore the possibility that the racial differences in bunching at 280g are driven by another factor correlated with race, and I show that racial differences exist even among observably similar offenders.[7] Next, I test whether the disparity could be the result of racial differences in costs to the prosecutor of charging a defendant with 280g. I provide evidence that the racial disparity is not related to racial differences in defense, to judge characteristics, or to other costs of developing a case. Last, I consider statistical versus taste-based discrimination from prosecutors.[8] I show that the racial disparity in bunching can be mostly explained by a measure of state-level racial animus based on Google search data developed by Stephens-Davidowitz (2014). Black and Hispanic offenders convicted in states with higher levels of racial animus are more likely to be bunched at 280g than white offenders convicted in those states. In states with lower levels of animus, offenders are equally likely to be bunched at 280g by race. Persistent racial differences after interacting race with observables, within-district variation in prosecutor-level bunching, and the correlation between the racial disparity in bunching and state-level racism all support a model of discrimination in which the disproportionate use of discretion is a result of prosecutor tastes. A more detailed model of statistical discrimination could incorporate those facts, and I cannot reject such a model. Regardless, neither form of discrimination is innocuous; statistical discrimination may be rooted in structural or taste-based discrimination, can be based on inaccurate beliefs (Bohren et. al. 2022), and is not legally permissible (Rehavi and Starr 2014).

Taken together, these results suggest a subset of federal prosecutors use their discretion to tag some defendants with drug amounts that will trigger mandatory minimums, and that they do this disproportionately for black and Hispanic defendants. Even more, the decrease in bunching after the Supreme Court tightens evidentiary standards in *Alleyne* suggests these cases are reliant on relatively weak evidence. Several additional analyses suggest this racial disparity can be attributed to taste-based discrimination.

Ultimately, the racial disparity in bunching at 280g has meaningful implications for the racial sentencing gap. Depending on the counterfactual sentence imputed for the affected offenders, bunching at 280g can account for up to 6 percent of the racial disparity in crack-cocaine sentences. A conservative estimate suggests that being bunched at 280g adds 1-2 years to an offender's sentence (Tuttle 2022). Multiple estimates suggest the cost of incarceration (combining direct care costs and the cost of lost current and

[7]Again, if racial differences were driven by these other factors, this would not imply that race or discrimination in general are unimportant. Those other differences could themselves be a product of discrimination and broader structural inequalities.

[8]While these models of discrimination have been the subject of enduring interest in the economics literature, other fields emphasize that they do not capture all forms of discrimination or racism (Small and Pager 2020).

future wages for the offender) is approximately $60,000 per person per year (Donohue 2009; Mueller-Smith 2015). I find 3.4% of black and Hispanic crack-cocaine offenders are bunched at 280g after 2010 versus 1.3% of white crack-cocaine offenders. Assuming 3.4% and 1.3% of all drug cases from 1999-2015 were subject to similar discretion by race implies total costs of 1.3 billion dollars for black and Hispanic offenders versus 173 million dollars for white offenders. In terms of incarceration, the disparity implies 21,000 years sentenced due to this discretion for black and Hispanic offenders versus 2,900 years for white offenders.

Broadly, this paper adds to an extensive literature in economics on racial disparities and discrimination in the criminal justice system, the majority of which focuses on police officers or judges. This paper also contributes to the empirical literature on prosecutorial discretion.[9] Prior work emphasizes the importance of prosecutor decisions for case outcomes. Mueller-Smith (2015) finds that prosecutors are among the strongest predictors of incarceration length in Harris County, Texas courts. Yang (2016) highlights the importance of prosecutor discretion in federal cases by studying charging and plea decisions in the face of court resource constraints. Agan, Doleac, and Harvey (2022) show that more lenient prosecutors can reduce recidivism by choosing not to prosecute certain cases. In this paper, I explore the discretion that prosecutors exercise and racial disparities in that discretion. Yang (2015) and Fischman and Schanzenbach (2012) document a racially disparate increase in sentence lengths concentrated at mandatory minimum lengths in federal court after judges are given more discretion over other aspects of sentencing.[10] In this paper, I focus on a change in sentencing rules rather than a change in judicial powers, allowing me to trace out the response to that change and identify a clear use of prosecutorial discretion. In a concurrent paper, Sloan (2022) uses random assignment of prosecutors to misdemeanor cases in New York County to show that being assigned to an opposite-race prosecutor increases a defendant's likelihood of conviction, particularly in property crime cases. In this paper, I focus on felonies in the federal system, documenting prosecutor-driven disparities in a large court system where defendants face particularly long sentences. Understanding disparities in the federal system is particularly important because it is larger than any individual state prison system, approximately 30% of people in prison for drug offenses are in federal prison, and federal sentences are typically more severe than state sentences (Wright 2006; PPI 2019).[11]

Understanding disparities and bias from prosecutors, particularly those in federal court or on felony cases, has proven difficult. Rehavi and Starr (2014) made a significant advance in this regard by introducing

[9] On disparities, see, for example: Knowles, Persico, and Todd 2001; Anwar and Fang 2006; Close and Mason 2006; Antonovics and Knight 2009; Abrams, Bertrand, and Mullainathan 2012; Anwar, Bayer, and Hjalmarsson 2012; Rehavi and Starr 2014; Park 2017; Pfaff 2017; Arnold, Dobbie, and Yang 2018; West 2018; Ba, Knox, Mummolo, and Rivera 2020; Cox and Cunningham 2020; Feigenberg and Miller 2020; Hoekstra and Sloan 2022; Luh 2022; Sloan 2022. On prosecutor discretion, see, for example: Glaeser, Kessler, and Piehl 2000; Bjerk 2005; Boylan 2005; Shermer and Johnson 2010; Fischman and Schanzenbach 2012; Rehavi and Starr 2014; Mueller-Smith 2015; Yang 2015; Yang 2016; Nyhan and Rehavi 2017; Silveira 2017; Arora 2018; Krumholz 2019; Didwania 2020; Harrington and Shafer 2020; Jordan 2021; Ouss and Stevenson 2022; Sloan 2022; Agan, Doleac, and Harvey 2022. Literature outside economics is discussed briefly in Section II.

[10] One explanation of these results is that judges are more constrained by mandatory minimum sentences once they are given more discretion over other aspects of sentencing.

[11] Studying the federal system also allows me to examine the response to a Supreme Court decision that changed evidentiary standards and allows me to analyze geographic variation in discretion across federal districts located in all US states.

linked data from arrest to sentencing in federal court that permitted a detailed selection on observables approach. They found that black offenders receive harsher sentences than white offenders arrested for the same crime, and that this disparity is driven by prosecutorial discretion to bring a charge with a mandatory minimum. A primary concern that they highlight is that even within observed offense codes, there may be unobserved differences in criminal conduct. Since their contribution, the literature on prosecutors, outside of a few papers that focus on state or local courts, has not made the same advancements as the larger literature on policing and judges. This is largely due to the fact that: (i) once a case reaches sentencing it involves a complex array of decisions and actors, (ii) the data offer limited information about those decisions and actors, and (iii) in these high stakes federal cases, cases are not randomly assigned to prosecutors.

In this paper, I contribute to the literature in several ways. First, I address challenges (i) and (ii) above by identifying a clear use of discretion–the bunching of cases at 280g–and pulling together data from drug use surveys, drug seizures, state court records, prosecutor case files, and final sentencing to get a complete picture of when and where this use of discretion arises. In doing so, I provide evidence on prosecutor discretion in federal court using this new source of identification–the sharp change in the crack-cocaine mandatory minimum threshold in 2010. I then show that the burden of this discretion falls disproportionately on black and Hispanic offenders, documenting how prosecutor discretion influences racial sentencing gaps. Second, I address challenge (iii) above by using the distribution of amounts charged pre-2010 as an empirical counterfactual to the post-2010 distribution. I document that the distributions were similar by race prior to 2010, indicating similar drug involvement by race. A key question when studying sentencing disparities is whether those disparities are driven by differences in criminal conduct or differences in treatment conditional on conduct. Drug amount is a key input for sentencing and is thus a more granular and highly relevant measure of criminal conduct and offense severity than has been used in prior studies. This empirical strategy provides new, compelling evidence that the disparate treatment of offenders in this setting is not due to differential drug involvement. Altogether, this constitutes new evidence on prosecutor discretion in federal court and in high stakes felony cases, two critical yet understudied settings. Third, I quantify the fraction of prosecutors exercising this discretion, and I show that it can be mitigated by increasing evidentiary standards. While prior work examines disparities in the use of discretion, the analysis of *Alleyne* offers a new test of the quality of evidence underlying the use of discretion. Finally, I provide evidence on why this disparity arises by ruling out many alternative mechanisms and documenting a robust correlation between the racial disparity in bunching and state-level racial animus.

Along with work by Anbarci and Lee (2014) and Goncalves and Mello (2021), I implement this new test for racial bias in criminal justice using insights from the bunching literature. Anbarci and Lee (2014) and Goncalves and Mello (2021) study police discounting speeding tickets and show that officers exhibit racial bias in their leniency. I contribute to this strand of the literature in multiple ways. First, I identify bunching from policy-induced variation in the punishment thresholds rather than cross-sectional variation in officer

propensity to bunch at a fixed threshold. Second, I identify this use of discretion and the actors responsible in the complex federal court setting where many more agents intertwine and where the resulting sentences are particularly high. Finally, I focus on prosecutors, a pivotal agent in the criminal justice system who is afforded wide latitude over numerous aspects of a case, including case dismissals, charging decisions, and plea offers. It is critical to understand how an agent who is given this vast discretion uses it in practice.

II. Institutional Background and Prosecutor Objectives

A. Institutional Background

1. The Fair Sentencing Act, Mandatory Minimums, and Drug Quantities

Debate about federal mandatory minimum policy has overwhelmingly focused on the disparity between the thresholds for crack- and powder-cocaine. Prior to 2010, the threshold for the crack-cocaine 10-year mandatory minimum was 50g whereas the 10-year threshold amount for powder-cocaine was 5000g, a 100-to-1 disparity. Academics, activists, and policymakers have long recognized the racially disparate consequences of the crack-powder disparity (e.g. Taifa 1998; Alexander 2010). In part due to the recommendations of the United States Sentencing Commission (USSC) and in part due to the political climate, the threshold amounts for crack-cocaine were increased in August 2010 by the Fair Sentencing Act (FSA). The upper threshold was changed from 50g to 280g, and offenders sentenced after the FSA were subject to the new threshold.[12] In this paper, I use this change from 50g to 280g to study bunching at mandatory minimum thresholds and its relation to discretion and racial disparities in the criminal justice system.[13]

This paper is not the first to acknowledge bunching in the amount of drugs recorded in US federal sentencing or its importance as a potential sign of discretion, but it is the first to use the time variation in the mandatory minimum threshold to test for a racial disparity conditional on drug involvement and then conduct a comprehensive analysis of where and why that disparity arises.[14] Bjerk (2017b) briefly discusses bunching in the distribution of drug amounts, but posits that bunching arises from negotiation

[12] It is not clear why 280g was chosen. One potential reason is that lawmakers wanted to set the threshold at 10 ounces (283.495g), but in keeping with the convention of setting it in grams or kilograms, chose 280g as the closest "round" number.

[13] One concern with limiting focus to bunching at 280g is that the racial disparity in bunching could be offset by a disproportionate shifting of white offenders to even greater amounts, such as 300g or 350g. Why focus on cases bunched narrowly above 280g when any amount above 280g will trigger the mandatory minimum? Conceptually, cases bunched at 280g are cases in which the prosecutor is charging the minimum amount necessary to trigger the mandatory minimum, suggesting that they may be based on weaker evidence than cases charged with amounts far above 280g. Empirically, there is a substantial increase in bunching at 280g, but only small increases above 290g after 2010. After accounting for long-run trends in drug weights charged, those small increases are the same magnitude for all offenders. There is no offsetting shift to even larger amounts for white offenders. I discuss these tests in more detail in Section V.A.2 and the Online Appendix. In Section V.A.3, I test whether black and Hispanic offenders are moved away from narrow ranges from 60-280g at the same rate as white offenders after 2010. If there is an offsetting shift for white offenders, we should see that black and Hispanic offenders are moved away from narrow ranges at similar rates as white offenders, but that black and Hispanic offenders are shifted to 280-290g while white offenders are shifted above 290g. In Section V.C, I explore how bunching changes after a Supreme Court case that tightens evidentiary standards to test whether the bunched cases are, in fact, based on weaker evidence. These tests confirm the conceptual motivation for studying bunching at 280g.

[14] In concurrent work, Knorre (2020) shows bunching in reported drug amounts from Russian police, and Travova (2020) studies how that bunching relates to performance evaluation incentives. Lepage (2020) finds that offenders strategically respond to changes in state-level cutoffs for lower-level drug possession, noting that the response only differs by race for marijuana.

downward by prosecutors and defendants.[15] A 2015 Bureau of Justice Statistics (BJS) working paper on federal sentencing disparities also investigates the idea that prosecutors could "game" the drug weight sentencing guidelines (Rhodes, Kling, Luallen, and Dyous 2015). That report provides a cursory look at bunching above mandatory minimum thresholds for all drugs by race, but does not address the bunching that is always present at round-number amounts (50g, 100g, 500g, etc.). As such, the authors conclude prosecutorial discretion in this form does not differentially affect black and Hispanic offenders.[16]

I depart from previous work in several ways. First, I show that excess mass at the threshold comes from cases below the threshold rather than above it. I also show that the bunching is more pronounced in trial cases, which suggests that drug amounts are being moved above the cutoff and not negotiated down to it. Second, I use the time variation in the crack-cocaine threshold to isolate bunching that is solely due to prosecutor choices. Finally, I examine data at multiple stages in the criminal justice process and conduct several empirical tests that all suggest prosecutorial discretion negatively affects minority defendants.

2. Procedural Background

In Figure A1, I illustrate a simplified timeline from arrest to sentencing. Arrests are made by local police or federal agents, and after arrest, cases are handled by state or federal prosecutors.[17] Prosecutors decide whether to prosecute or decline the case. Federal arrests typically stay in the federal system, but local arrests can be shifted to federal court or tried in both state and federal court (Wright 2006). A case prosecuted in federal court can end in conviction, acquittal, or dismissal. For convictions, a probation officer, partly in consultation with the prosecutor, prepares a pre-sentence report (PSR) that details facts relevant to sentencing. At sentencing, the judge considers statements from the prosecution, the defense, and the PSR to make factual determinations (e.g. the amount of drugs involved) and decide the sentence. In 2015, approximately 70% of drug arrests referred to federal prosecutors were prosecuted and 90% of those prosecuted ended in a conviction (BJS 2019). The drug quantity used in sentencing can be influenced at many of these stages. Below, I describe the legal discretion that prosecutors have over the drug quantity.

Prosecutors can influence drug amounts because mandatory minimum sentencing is determined by the total amount of drugs the offender is responsible for trafficking, which is not strictly based on the amount of drugs they are holding at the time of arrest (Honold 2014; USSC 2015b; Lynch 2016). For one, prosecutors

[15] Since Bjerk focuses on consequences of mandatory minimums for all drug types, he does not investigate the cause of the observed bunching in crack-cocaine offenses. In addition, that paper does not explore differences before and after the FSA

[16] Moreover, the working paper does not trace out the distribution of crack-cocaine amounts across stages of the criminal justice system, and as such, does not pinpoint where bunching in crack-cocaine first occurs. Also, it does not estimate racial differences in the distribution of amounts, and thus does not explore whether the differences in bunching arise from differences in drug involvement or whether the racial differences are conditional on drug involvement. Finally, it does not estimate the effect of changing evidentiary standards on the level of bunching and does not explore why disparities exist in bunching of crack-cocaine amounts. The working paper is an extensive and excellent treatment of sentencing disparities more broadly. In that light, it is reasonable that the authors did not do a "deep dive" on this specific test, which is a small piece of the broader report. I do a "deep dive" in this paper and come to substantively different conclusions, introduce numerous new findings, and ultimately, provide new evidence on prosecutor discretion.

[17] Throughout, I often use the word "police" to refer broadly to local police and federal agents.

can rely on the testimony of informants or law enforcement to establish the amount of drugs a defendant is responsible for outside of the actual drugs seized (Lynch 2016). In addition, mandatory minimums also apply to trafficking conspiracies in which the total amount trafficked by the group in question can be applied to all members of the group (Lynch 2016). The USSC Guidelines (2015b) state, "Where there is no drug seizure or the amount seized does not reflect the scale of the offense, the court shall approximate the quantity of the substance." Online Appendix B of Tuttle (2022) contains further details and examples.

In Section V.B, I examine data from several stages to estimate the source of the bunching at 280g. Sections V.B-V.D further rule out explanations related to offenders, police, defense attorneys, probation officers, and judges. Ultimately, I find that prosecutorial discretion leads to the bunching at 280g.

B. Prosecutor Objectives and Responses to the Fair Sentencing Act

The fields of economics, criminology, and law all admit self-interested and/or biased prosecutors.[18] The canonical economic model of the courts from Landes (1971) assumes that prosecutors maximize the expected sum of sentences subject to resource constraints. Others have modeled resource-constrained prosecutors trying to achieve an ideal punishment for guilty parties and no punishment for innocent parties (Grossman and Katz 1983; Reiganum 1988), yet empirical work finds that prosecutors are, in part, career-focused (Glaeser, Kessler, and Piehl 2000; Boylan 2005). In addition, recent work shows racial bias in prosecutorial decisions (Rehavi and Starr 2014; Sloan 2022).

Discussions of prosecutorial discretion in law reviews frequently note that career-oriented prosecutors focus on securing lengthy sentences or high conviction rates (e.g. Bibas 2004; Simon 2007; Barkow 2010; Sklansky 2018). Stuntz (2004) argues that prosecutors lean on harsh sentences to secure guilty pleas and that sentencing guidelines allow them to make credible threats. Criminologists, political scientists, legal scholars, and sociologists have studied prosecutorial bias along race, gender, and partisan lines (e.g. Spohn, Gruhl, and Welch 1987; Mustard 2001; Farrell 2003; Ulmer, Kurlychek, and Kramer 2007; Gordon 2009; Shermer and Johnson 2010; Davis 2013; Ulmer, Painter-Davis, and Tinik 2014; Franklin and Henry 2020).[19]

This work from economics, criminology, and law suggests that prosecutors may value crossing the mandatory minimum threshold in drug cases (for sentence length or certainty) and that they may value it differentially by race (due to racial bias). Prosecutors may desire high sentences due to career concerns, beliefs that long sentences are ideal (for retribution or future deterrence), or to wield them as tools in plea bargaining. By law, cases above the threshold must receive a sentence of at least five or ten years unless specific conditions are met for a sentence departure.[20] In practice, longer sentences are handed down in cases just above the threshold. Online Appendix C of Tuttle (2022) estimates the sentencing consequences

[18] Although this conceptual discussion describes prosecutor objectives as homogenous, I ultimately find that only a subset of prosecutors behave in this way in relation to bunching at the mandatory minimum threshold.

[19] Light (2022) notes raw disparities in federal sentencing have nearly disappeared but highlights, "the extent to which black-white inequality changed in recent years in the highly discretionary arrest, charging, and conviction stages awaits further inquiry."

[20] Note, these departures do reduce sentence length when applied, but to be eligible for these departures, defendants must cooperate with the government. I detail the costs of cooperation in the notes of Table A6.

of bunching at 280g. Assuming that introducing new evidence is costly and the cost is increasing in the amount of new evidence introduced, these objectives yield predictions for how prosecutors will behave in the face of thresholds and how they will behave when those thresholds change.[21]

Prior to 2010, the mandatory minimum thresholds in federal court for crack-cocaine were 5g (for a 5-year minimum) and 50g (for a 10-year minimum). After 2010, these thresholds shift to 28g and 280g. This shift should lead to the following relative changes: (1) an increase in the density from 0-5g, (2) an ambiguous change from 5-28g, (3) an increase from 28-50g, (4) a decrease from 50-280g, (5) an increase from 280-290g, and (6) no change above 290g.[22] See Figure A2 for an illustration of these changes.

These changes should occur because some cases worth bunching at 5g or 50g before 2010 will also be worth bunching at 28g or 280g after 2010 and some will no longer be worth it.[23] Also, some cases that were not bunched before 2010 will be worth bunching after 2010. Online Appendix D of Tuttle (2022) discusses these changes in the context of a simple model of prosecutor objectives. In Section V.A.2, I show empirical evidence consistent with this conceptual discussion.

In 1983, legal scholar and eventual judge Frank Easterbrook wrote, "Rules could command, for example, that all cases involving a sale of cocaine weighing more than 50 grams be prosecuted and all others not. Rules of this sort produce the arbitrary and unexpected consequences so well known to tax and welfare lawyers; it is far from clear that one can design rules to achieve a particular end. People will change their conduct to take advantage of lacunae." Since then, such rules have been implemented, but researchers have paid scant attention to the ways people have changed their conduct to take advantage of them. In this paper, I document changing conduct by prosecutors that disproportionately affects black and Hispanic defendants–behavior that has been discussed and researched qualitatively by legal scholars and criminologists but that has remained relatively unexplored empirically.

III. Data

To estimate the degree of bunching at the 10-year mandatory minimum threshold, I use data on federal cases that include the amount of drugs recorded at sentencing. I then bring in several other datasets from different stages in the criminal justice process to estimate who is responsible for the bunching at 280g.[24]

Figure A1 shows a simplified timeline from arrest to sentencing and describes how the data I use is related to each step. This timeline acknowledges that selection into the data can occur at each step. As Knox, Lowe, and Mummolo (2020) discuss, bias in selection into the dataset of interest can distort the

[21] The cost of introducing new evidence, especially as that evidence grows weaker, can include costs of building the case, reputational costs with regard to colleagues and judges, psychological costs, the risk that the judge or jury will not make the same factual determination, and costs associated with agreeing on the amount in the plea bargaining process.

[22] Note, whenever ranges are listed, the upper bound of the range is not inclusive.

[23] A case is worth bunching if the benefit (e.g., increased sentence) exceeds the cost (e.g., costs of introducing new evidence)

[24] I am not able to link people across these datasets. However, given the nature of the findings and the information available in each dataset, analyzing them independently is sufficient to show where the bunching first occurs and to rule out alternatives. Linked data from arrest to sentencing does exist, but the codebook suggests it lacks a measure of quantity seized at arrest.

ultimate measure of bias. My empirical approach takes any bias in selection as given, and assumes this bias does not change sharply in 2010. I show evidence of this: drug selling and crack-cocaine use does not increase after 2010, quantities seized do not increase after 2010, and the composition of cocaine offenses in state convictions does not change after 2010. Penalties remain high for offenses with less than 280g, suggesting there is little reason for selection into federal sentencing to change pre- versus post-2010.[25] Also, Rehavi and Starr (2012) use linked data to show that the probability a case is filed in federal court and the probability a defendant is convicted is the same for black and white defendants (conditional on arrest). Finally, as long as selection into the data is biased in favor of white defendants (e.g. police are more lenient with white defendants or prosecutors are more likely to dismiss white defendant cases), then the estimate of the racial bias in this paper will be an underestimate.

A. United States Sentencing Commission (USSC) Data

To estimate the degree of bunching at or above 280g, I use data provided by the USSC on recorded drug amounts in all federal drug cases sentenced from 1999-2015.[26] I focus on cases that involve a crack-cocaine offense since that is the only drug for which the mandatory minimum threshold changes over time. Approximately 7.8% of offenders in this sample are labeled as white, 10.6% as Hispanic, and 81.6% as black. Table 1 summarizes additional offender and offense characteristics.

I restrict these data to cases in which the amount of drugs is non-missing and is not recorded as a range. Approximately 20% of cases are excluded because the weight is recorded as a range and 2% are excluded because the weight is missing. The fraction of cases excluded for these reasons does not change discontinuously at 2010, though it does increase in 2013 and 2014. In the Online Appendix, I show that including cases coded as a range only exacerbates the degree of bunching and the racial disparity in bunching. I also remove cases that are flagged for having data issues with the drug quantity variable, the sentencing variable, or cases where the court does not accept or changes the findings of fact. In general, I limit the sample to 0-1000g to remove outliers, but results are robust to removing this restriction. Online Appendix E of Tuttle (2022) is a data appendix which discusses this dataset, data construction, other minor restrictions, and robustness of the main results to these choices in more detail.

B. Additional Data

In addition to the USSC data, I incorporate several other datasets to understand the source of the bunching in drug cases. I describe the six main datasets here.

Florida State Inmate Database, 2000-2015. These data include the year of conviction, a description

[25]The aggregate Federal Justice Statistics do not exhibit a sharp break in case handling after 2010. From (Oct.) 2009-(Oct.) 2010, the DEA made 2,408 arrests of black or Hispanic people for crack-cocaine offenses and 451 arrests of white people compared to 2,448 and 496 arrests from 2011-2012. The Office of the US Attorney prosecuted 75.1% of drug cases received from '09-'10 and 75% from '11-'12. Among prosecuted cases, 93.1% ended in a conviction from '09-'10 compared to 92.6% from '11-'12.

[26]These amounts are derived from pre-sentence reports prepared by a probation officer and in consultation with the defense and prosecution. In the event the court rejects an amount in the pre-sentence report, the new amount is recorded in the statement of reasons report and reported in the USSC drug quantity field.

of the offense, and the offender's race. In Florida, offense descriptions typically include the name of the drug involved, and occasionally, a range for the amount of drugs involved (these broad ranges are: 0-28g, 28-200g, 200-400g, and 400+g). Florida does not separately categorize crack versus non-crack cocaine offenses, instead labeling all such offenses as "cocaine." Summary statistics for these data, the NIBRS drug seizures, and the DEA drug exhibits are reported in Table A1a.

National Incident Based Reporting System (NIBRS) Property Segment, 2000-2015. The FBI collects incident-level reports from local law enforcement agencies to create the NIBRS dataset. These reports are submitted voluntarily by agencies and thus, are not representative of national or state-level crime. For this reason, I use a balanced panel of agencies from 2000-2015. The property segment of this database includes information about drug seizures and drugs involved in arrests.[27] The offender segment of this database includes information on offender race, sex, and age for all offenders involved in the incident.

DEA System to Retrieve Information from Drug Evidence (STRIDE), 2000-2015. The STRIDE database contains information about drug evidence from the DEA and other agencies that was submitted to DEA laboratories for analysis. I obtained the data from a Freedom of Information Act request for records pertaining to the drug "cocaine" from 2000 to 2015. This information includes year and month the drugs were acquired, the price from undercover purchases, and drug weight, type, and potency.

Executive Office of the US Attorney (EOUSA), Caseload Data, 2000-2017. The EOUSA releases case-level data on cases processed by the US Attorney's office. These data are derived from information entered into the Legal Information Office Network System (LIONS) case management system. The EOUSA notes that each district may use LIONS differently, and as such, the data should not be used to make cross-district comparisons. The analyses using these data are robust to the inclusion of district fixed effects and various methods of accounting for missingness in the drug quantity data (a data quality issue that varies across districts). The EOUSA data includes information on type of drug, recorded quantity, an ID for the lead attorney on the case, and an ID for the judge on the case. Summary statistics are reported in Table A1b.

National Survey on Drug Use and Health (NSDUH), 2002-2016. The NSDUH is a survey of non-institutionalized US civilians aged 13 or older that primarily asks questions about drug use and mental health. The respondents are randomly sampled based on state and age, with larger states and younger individuals oversampled. I use two questions asked from 2002-2016: (1) "have you ever, even once, used crack-cocaine?" and (2) "during the past 12 months, how many times have you sold illegal drugs?" These data provide detail about drug use and drug selling that is not based on interactions with law enforcement.

Google Search Trends Data on Racial Animus from Stephens-Davidowitz (2014), 2004-2007. To measure racial animus at the state-level, I use data introduced by Stephens-Davidowitz (2014). Stephens-Davidowitz uses Google search data from 2004-2007 (accessed via the Google Trends tool) and measures

[27] Upon receipt of the data, the FBI checks for errors and contacts agencies for corrections. To the best of my knowledge, there are no known issues with the drug quantity field of the property segment. However, see Bibel (2015) for a discussion of well-known issues with NIBRS data more broadly, such as reporting and measurement of sexual assault, differential coverage, etc.

relative search volume in every US state for a specific racial slur and its plural form. Since Google searches are virtually anonymous, this measure may provide a less filtered view of racial attitudes than common survey measures. It is positively correlated with racial animus as measured by implicit association tests or questions about interracial marriage from the General Social Survey.[28] The construction of the measure is covered in much greater detail in Stephens-Davidowitz (2014).

IV. Methodology

This paper has four main goals. First, to quantify the bunching at 280g after 2010 and the racial disparity in bunching at 280g. Second, to estimate whether the racial disparity in bunching at 280g is due to differences in the underlying distributions of drug involvement or a difference in the likelihood a case is bunched **conditional** on drug involvement of the defendant (i.e. a **conditional racial disparity**). Third, to estimate who causes the bunching at 280g after 2010. And fourth, to test various explanations for the racial disparity in bunching, including discrimination. In this section, I detail methodology for the first three goals. I reserve the discussion of discrimination and related tests for Section V.D.

Throughout, I use what Kleven (2016) terms the "difference-in-bunching" method. This approach estimates the degree of bunching by comparing the actual distribution to an empirical counterfactual distribution. To estimate bunching at 280g and the racial disparity in bunching, the ideal counterfactual is the post-2010 distribution with the pre-2010 thresholds. I assume the pre-2010 distribution is a good counterfactual in this sense for all parts of the drug quantity distribution. Section IV.A details the estimation of bunching and the racial disparity under this assumption.

To estimate a conditional racial disparity in bunching at 280g, the ideal counterfactual is the post-2010 distribution with no mandatory minimum threshold (or other incentive to increase the amount charged). I assume the pre-2010 distribution is a good counterfactual in this sense for the part of the quantity distribution above 50g. Section IV.B outlines tests for a conditional racial disparity under this assumption and discusses the implications of potential bias in the pre-2010 distributions.

Finally, to estimate who causes the bunching at 280g, I test for changes in drug quantity at multiple stages in the criminal justice process leading up to sentencing. Here, again, the assumption is that at each of these stages the pre-2010 distribution is what the post-2010 distribution would be if the thresholds had not changed. Thus, I use the same methods detailed in Section IV.A. In the Results section, I detail methodology for several additional analyses.

A. Bunching at 280g and Racial Disparity in Bunching

I define a case as "bunched" at 280g as any case in the narrow range 280-290g (not including 290g). I then compare the fraction of cases from 280-290g in the post-2010 distribution of drug weights to the fraction of

[28] It is also correlated at the Census region level with responses to these questions from respondents with a graduate degree. This suggests it is not solely reflective of racism from people with low levels of formal education.

cases from 280-290g in the pre-2010 distribution. I estimate the following linear probability model:

$$(\text{Charged } 280-290\text{g})_{it} = \alpha + \beta \text{After2010}_{it} + \epsilon_{it} \tag{1}$$

where $(\text{Charged } 280-290g)_{it}$ is equal to one if offender **i** in year **t** is charged with 280-290g and is equal to zero if the offender is charged with less than 280g or equal to or above 290g.[29,30] After2010_{it} is equal to one if the offender **i** in year **t** is sentenced based on the Guidelines amendment years 2011-2015 and is equal to zero if the offender is sentenced based on the Guidelines amendment years 1999-2010. β is the change in an offender's probability of being charged with an amount in the narrow 280-290g range as a result of being sentenced after the threshold amount is increased to 280g. To estimate heterogeneity in bunching by race, I extend the model as follows:

$$\begin{aligned}(\text{Charged } 280-290\text{g})_{it} = \alpha &+ \beta(\text{After2010} \times \text{White})_{it} \\ &+ \delta(\text{After2010} \times \text{BlackOrHispanic})_{it} + \text{BlackOrHispanic}_{it} + \epsilon_{it}\end{aligned} \tag{2}$$

Now, β represents the change in a white offender's probability of being charged with 280-290g as a result of being sentenced after 2010, and δ represents the change for black and Hispanic offenders.[31,32] I include these terms separately to show the change in bunching directly for each group, and I provide the p-value from the test that these coefficients are equal. To understand where the excess mass at 280-290g comes from (i.e. where the post-2010 distribution has less mass relative to the pre-2010 distribution), I estimate models similar to equation (1) that replace the dependent variable with different quantity ranges:

$$(\text{Charged X-Yg})_{it} = \alpha + \beta^X \text{After2010}_{it} + \epsilon_{it} \tag{3}$$

Here, β^X is the change in an offender's probability of being charged with an amount of drugs between X and Y grams as a result of being sentenced after the threshold is increased. I estimate equation (3) for 0-5g, 5-28g, 28-50g, 50-60g, 60-100g, 100-280g, 280-290g, 290-470g, 470-600g, and 600-1000g.

B. Racial Disparity Conditional on Drug Involvement

Now, I outline two tests to estimate whether the racial disparity in bunching at 280g is due to differences in the underlying distributions of drug involvement by race or a difference in the likelihood a case is bunched

[29] State conviction data does not include precise drug weights. In those cases, I use the dependent variable (Convicted with 200-400g), equal to one if the offender is convicted with 200-400g and equal to zero otherwise.

[30] I quantify the excess mass at 280-290g by using regression analysis on the case-level microdata. This follows work by: Kleven et al. (2011), Traxler et al. (2018), Goncalves and Mello (2021). I am primarily interested in estimating the change in the probability a case is charged with 280-290g after 2010 and whether that change differs by race. Online Appendix F of Tuttle (2022) shows that the results in this paper are robust to alternative methods of quantifying bunching.

[31] Combining black and Hispanic offenders into one category, although common in analyses of the criminal justice system, is a crude categorization. In this analysis, splitting these groups yields similar results. There is a larger increase in bunching for black offenders than white offenders and a larger increase for Hispanic offenders than white offenders. The increase in bunching is similar for black and Hispanic offenders. See Figure A3e where I differentiate these defendants. For expositional reasons, I combine these groups throughout the paper. However, it is worth emphasizing that their experience with law enforcement and with discrimination in the US, in general, is varied and complex in a way that is not accounted for in this analysis (RWJF 2018).

[32] Black and Hispanic offenders are slightly more likely (0.2 percentage points) to be charged in the 280-290g range prior to 2010. This approach addresses the small pre-existing difference by estimating the increase in probability after 2010 by race.

conditional on drug involvement. Online Appendix G of Tuttle (2022) discusses a model motivating these tests. Showing the disparity exists conditional on drug involvement does not necessarily imply the disparity is unrelated to other aspects of the case. Section V.D tests various explanations for the racial disparity.

For the first test, consider that cases bunched at 280g are primarily cases that would have been bunched at 50g prior to 2010 or charged with an amount from 60-280g prior to 2010.[33] If the racial disparity in bunching at 280g is explained by racial differences in movement away from the 50g threshold, then it will be impossible to say (with the available data) whether the disparity in bunching at 280g is a conditional racial disparity. This is because racial differences in the shift away from 50g could be related to differences in the underlying distribution of drug involvement that initially caused the disproportionate bunching at 50g. To test whether the racial gap in bunching at 280g can be explained by racial differences in shifting away from the 50g threshold, I estimate equation (4):

$$\begin{aligned}(\text{Charged } 50-60\text{g})_{it} = \alpha + \beta^{50}(\text{After2010} \times \text{White})_{it} \\ + \delta^{50}(\text{After2010} \times \text{BlackOrHispanic})_{it} + \text{BlackOrHispanic}_{it} + \epsilon_{it}\end{aligned} \tag{4}$$

I add δ^{50} from equation (4) to δ from equation (2) to capture the disparity in bunching that remains after accounting for racial differences in shifting away from the 50g threshold.

The racial disparity that remains can be explained by: (1) racial differences in the distribution of cases from 60-280g prior to 2010 or (2) racial differences in movement away from specific ranges from 60-280g after 2010. For example, if black and Hispanic offenders are more likely to be charged with high amounts (e.g. 260-270g) prior to 2010, then the bunching could arise because prosecutors are more likely to shift those high weight cases to the new 280g threshold. I test whether there is a difference in the distribution of amounts charged from 60-280g prior to 2010 with a Kolmogorov-Smirnov test for equality of distributions. Under the assumption that the pre-2010 amounts from 60-280g represent true drug involvement, then equal distributions by race implies similar involvement. If the amounts are biased against black and Hispanic defendants pre-2010, then equal distributions by race implies lower drug involvement for black and Hispanic defendants. Equal distributions by race would only imply higher drug involvement for black and Hispanic defendants if the charged amounts were biased against white defendants prior to 2010.

For the second test, I directly test for a racial difference in shifting away from those narrow ranges from 60-280g. Specifically, I estimate the following:

$$(\text{Charged X-Yg})_{it} = \alpha + \delta^{X}(\text{After2010}\times\text{BlackOrHispanic})_{it} + \gamma\text{After2010}_{it} + \lambda\text{BlackOrHispanic}_{i} + X_{it} + \epsilon_{it} \tag{5}$$

where the ranges considered are 60-70g, 70-80g,...,260-270g, and 270-280g. Then, $\delta^{X} < 0$ implies that black and Hispanic offenders are more likely to be shifted away from a given amount X after 2010. Again, if black and Hispanic offenders are more likely to shifted away from a narrow range than white offenders,

[33] Cases bunched at 280g could be cases that would be recorded below 50g pre-2010 if there are fixed costs of bunching a case. Table 3 suggests that bunching at 280g primarily comes from cases that would have been recorded with 50-280g pre-2010.

that suggests the racial disparity in bunching at 280g is conditional on drug involvement.

V. Results

A. Main Results

1. Primary Bunching Estimates and Robustness

Using the USSC data, I estimate the effect of being sentenced after 2010 on whether an offender is sentenced for an amount between 280-290g. Column 1 of Table 2 shows that offenders sentenced after the threshold increases to 280g are 3.3 percentage points more likely to be charged with amounts just above 280g. Column 2 shows that this increase in bunching is driven by black and Hispanic offenders, who are approximately three times as likely to be charged with 280-290g after 2010 compared to white offenders. Figures 1a-d display graphical evidence of bunching at 280-290g and the racial disparity in that bunching.[34]

This result is robust to various sample restrictions; the inclusion of state fixed effects, time trends, state-specific time trends, and offender-level controls; clustering standard errors at the state-level; the use of Logit/Probit/Poisson models; wider definitions of the bunching range (e.g. 280-380g); the inclusion of cases with weights coded as range; and alternative methods of calculating standard errors. See Tables A3-A4 for these results. The results are also robust to re-calculating standard errors using randomization inference.[35] In addition to these tests, approximately 70% of the disparity remains after controlling for state-by-post fixed effects and 60% remains with district-by-post fixed effects. The disparity is also robust to removing one state or one district at a time from the sample. I also conduct a simple bounding exercise in Table A5 that accounts for potential substitution into other drug types or selection into the case's drug weight being coded as a range. Table A5 also presents a difference-in-differences analysis of bunching using other drug types for which the threshold did not change. Table A6 examines the use of safety valve and substantial assistance departures. These additional tests confirm the main results. Offenders sentenced after 2010 are more likely to be charged with 280-290g, and this increase is disproportionately large for black and Hispanic offenders.

2. Source of the Excess Mass at 280g

To understand the reason for this bunching at 280g, I analyze other parts of the quantity distribution. If the excess mass comes from above 290g, bunching may be the result of negotiation between prosecutors and defendants (Bjerk 2017b). However, if it comes from below 280g, prosecutors may be charging amounts

[34] Figures 1a-b zoom in on the most relevant part of the quantity distribution. See Figures A3a-b for a plot of the histograms from 0-500g. Figures A3c-f present alternative ways to visualize this phenomenon. In particular, Figure A3f shows that the total number of cases at 280-290g increases after 2010.

[35] Table 1 indicates there are only 544 white offenders in the data after 2010. One concern is that the disparity in bunching has occurred by chance. Traditional null hypothesis testing suggests this is not the case, but to further assuage concerns, I conduct an alternative test. I run 500 placebo regressions, randomly assigning white offenders to the 280-290g range at the same rate as black and Hispanic offenders pre- and post-2010. I only estimate a disparity as large as the actual disparity in two out of 500 cases. The median of these placebo estimates is 0.00008, the 90th percentile is 0.011, and the true disparity is 0.021.

above the threshold to secure longer or more certain sentences.[36]

In Table 3, I show the change in the probability of being recorded in different ranges. The estimated changes match the conceptual discussion in Section II.B. Summing the coefficients in columns 4-6 implies that the change in probability from 50g-280g can account for 80% of the increase in the 280-290g bin.[37] In Figures A4a-i, I plot the share of cases over time in each of these ranges. I estimate the regressions in Table 3 by race in Table A7. The results are similar but noisier since it requires cutting the already narrow ranges by race.[38] Figures A4j-k show results by race using broader ranges. Table A7 also shows that after accounting for long-run trends in these broad ranges, there is no offsetting shift to the 290-1000g range for white offenders. The results are noisier, but the magnitude of the racial disparity in whether a case is charged above 280g **at all** is as large or greater than the disparity in whether a case is at 280-290g.

Finally, I examine the degree of bunching in the subset of cases that go to trial. If the bunching is a result of plea bargaining or lenient prosecutors rounding down, we should expect less bunching in trial cases where the plea bargaining channel is shut down and incentives for leniency are muted.[39] However, the degree of bunching and the racial disparity in bunching is only heightened in trial cases (see column 3 of Table 2). In fact, the only cases with 280-290g that go to trial are those of black and Hispanic offenders. As before, the increased bunching is accompanied by a falling share of cases below 280g ($\beta = -0.107$ and $SE = 0.022$) and a small, rising share of cases above 290g ($\beta = 0.039$ and $SE = 0.020$). To explore robustness, I have also re-estimated the missing mass results in Table 3 using trial cases only and find similar results. In Section V.B.3, I show additional evidence from prosecutor case management data that cases bunched at 280g would likely be recorded below 280g in the absence of strategic prosecutor behavior around the threshold.

3. Estimating the Conditional Racial Disparity in Bunching at 280g

The results above show that there is a racial disparity in bunching at 280g. This alone is of interest–it shows that prosecutors use their discretion to increase sentences in response to the FSA and that the burden of

[36] It is impossible to say with certainty that the "missing mass" is where cases in the "excess mass" would be recorded had they been sentenced prior to 2010. As is typical in bunching analyses, I assume the missing mass is indicative of where the "excess" cases would be in the counterfactual. The empirical result is another piece of evidence that the bunching is a result of cases being shifted in a way that is consistent with a simple model of prosecutor behavior and the empirical evidence of no offender response.

[37] It is possible that some offenders charged with 280-290g post-2010 would have been charged below 50g pre-2010 if there are fixed costs associated with bunching (e.g. a fixed cost of evidence-gathering).

[38] Table A7 shows that white offenders and black and Hispanic offenders see similar increases in the probability they are charged with 28-50g after 2010. Because 28g is below the pre-2010 50g threshold, it is not possible to determine whether this is due to: shifting from below 28g, shifting from above 50g, or different drug involvement. Figure A4c, however, shows that there are long-run trends in the 28-50g range. Estimating a trend break model to account for these differential trends indicates that black and Hispanic offenders are 2.7 percentage points more likely to be charged with 28-50g after 2010 and white offenders are 1.1 percentage points more likely. Put more transparently, 11.1% of black and Hispanic offenders are charged with 28-50g in 2010 compared to 10.8% of white offenders. In 2011, the percentages increase to 13.5% and 11.6%, respectively. Given the concerns outlined above and in footnote 4 and the sensitivity to functional form, it's not clear how to interpret that difference.

[39] Prosecutors may still have an incentive to round down if they believe it will affect their probability of winning at trial. However, prior to the decision in *Alleyne* in 2013 the quantity determination is decided by the judge separately from the jury's verdict. In these cases, prosecutors should not expect the amount charged to relate to their probability of winning the case. I also find bunching at 280g and a racial disparity in bunching at 280g in trial cases decided before 2013.

this discretion falls on black and Hispanic defendants. However, this could be due to different underlying distributions of drug involvement by race or different treatment conditional on drug involvement. To disentangle these explanations, I conduct two tests outlined in Section IV.B.

The first test relies on decomposing the bunching at 280g. For the first part of that test, I estimate the racial difference in the shift away from 50-60g. I find that black and Hispanic offenders are less likely to be charged with 50-60g after 2010 ($\beta = -0.0054$, $se = 0.0030$). However, the decrease in the 50-60g range is not large enough to explain the disparity in bunching at 280g.[40] Adding the decrease from 50-60g for black and Hispanic offenders to the increase in 280-290g in column (2) of Table 2 yields a new coefficient of 0.0290. That coefficient is about 2.5 times larger than the coefficient for white offenders, and it is statistically different from the coefficient for white offenders at the five percent level (p-value = 0.0191).

For the second part of the first test, I test whether the distributions of charged amounts from 60-280g are equal by race prior to 2010. Figure 2a plots the distributions by race, and they are very similar. A Kolmogorov-Smirnov test of equality fails to reject the null that the distributions are equal (p-value = 0.792).[41] I argue that the pre-2010 distributions from 60-280g reflect true drug involvement because 50g is the highest mandatory minimum threshold at that time.[42,43] However, if the pre-2010 amounts are biased against black and Hispanic offenders, then the equal distributions imply that they have even lower drug involvement than white offenders. This test only implies higher drug involvement for black and Hispanic offenders if the pre-2010 amounts are biased against white offenders. Bias against white offenders only in the pre-2010 period would be in contrast to the existing literature on racial bias in criminal justice and would imply discordant directions of bias from prosecutors pre-2010 versus post-2010. Even more, alternative evidence from drug seizure records confirms black and white offenders are seized with similar amounts (see Panel A of Table 5 and Figures A7a-b) and a survey of inmates confirms drug involvement does not differ by race (see Tables A15a-b). Since the racial disparity in bunching at 280g cannot be accounted for by racial differences in movement from 50g or by racial differences in the distribution from 60-280g, this implies the disparity is a conditional racial disparity.

The second test for a conditional racial disparity in bunching relies on estimating racial differences in movement away from other narrow ranges. Figure 2b plots the coefficients from equation **(5)** divided by the share of cases in each range to show a percent difference by race.[44] There is a noisy decrease from

[40] White offenders also have a small decrease in the 50-60g range after 2010 ($\beta = -0.0036$, $se = 0.0114$).

[41] See Figures A5-A6 for these distributions broken out by criminal history, predicted sentence, and state-level racial animus.

[42] Sentences do increase slightly above 50g, but since there are no thresholds, the return to charging extra is much lower from 60-280g pre-2010 than post-2010. For example, the sentence increase from 150-160g to 280-290g post-2010 is over 7.5 times larger post-2010 than it is pre-2010.

[43] There are thresholds above 50g that correspond to different Sentencing Guidelines, however, the Guidelines are advisory after 2005, and I do not observe bunching at the Guidelines thresholds or a change in mass around those amounts once the Guidelines thresholds change in 2010. This is consistent with the Guidelines being advisory and prosecutors having less of an incentive to bunch at those amounts.

[44] Figure A7c shows estimates using 20g bins. In the main results, I only include estimates up to 100g above 280g because results are noisy at higher amounts. Figure A7d shows estimates up to 1000g. Figures A7e-g show additional robustness tests.

160-280g, but at several amounts, the coefficient is significantly different from zero or marginally significant. Further, the racial disparity in the overall decrease from 160-280g is not noisy, it is jointly significant at the 1 percent level. This implies that at those amounts, black and Hispanic offenders are more likely to be shifted away and bunched at 280g than white offenders. Again, this also implies a conditional racial disparity.

These results also provide evidence against an alternative explanation for the bunching at 280g. A common concern is that the cases bunched at 280g are cases that involve significantly more than 280g but are only charged with 280g because that is all that is necessary to trigger the mandatory minimum.[45] If prosecutors only record amounts necessary to trigger the minimum, the increase in bunching at 280g should come from a decrease in bunching at 50g. Also, the differential movement away from narrow ranges, such as 260-270g, after 2010 suggests that, despite basis for charging them similarly before, black and Hispanic defendants are treated differently after 2010. Even if the cases for black and Hispanic defendants truly involve over 280g and bunching is a result of disparate lenience for white defendants, these results raise the question: why aren't similarly positioned white defendants bunched at 280g at the same rate after 2010? Finally, as I show in section V.C, bunching at 280g decreases after a Supreme Court decision that tightens evidentiary standards.[46] If these cases bunched at 280g are cases where significantly more than 280g is involved, there is little reason for them to decrease in prevalence once evidentiary standards are raised.

B. Potential Mechanisms

1. Offender Behavior

If black and Hispanic offenders respond differently than white offenders to the FSA, a racial disparity in bunching at 280g may reflect prosecutors' reactions to those different responses. In Panel A of Table 5, I show that black and Hispanic offenders are not arrested with more drugs following the FSA, but instead, are holding slightly smaller amounts when arrested after 2010 (after controlling for state fixed effects, sex, and age).[47] In Panel B of Table 5, I show that black and Hispanic respondents to the NSDUH are not more likely to report having ever used crack or selling drugs in the past 12 months. This implies that the racial disparity in bunching cannot be attributed to differential responses in drug-carrying by race.[48]

2. Shifting of Cases Between State and Federal Courts

Drug Convictions in Florida Courts. The USSC data covers the universe of federal drug cases, but cases can be selected into federal court for many reasons. State and local authorities could send more of their high weight cases to federal court after 2010. Similarly, federal prosecutors could pull more of these types of cases from state and local courts after 2010.

[45] Cases are regularly charged with more than 280g and sentences increase slightly in amount charged above 280g.

[46] I also show in Table A8 that bunching at 280-290g in a district is correlated with the whether the district has any case reversed or remanded on appeal and particularly, whether it is reversed or remanded with an issue with the drug quantity.

[47] Likewise, I find no evidence of an offender response in the DEA STRIDE data on drug amounts or drug prices.

[48] Other papers also find small responses to changes in punishment. For example, Lee and McCrary (2017) find that people do not decrease offending at age 18, despite a discontinuous increase in the probability of a harsh sentence at that age.

To test this possibility, I use state-level data on cocaine offense convictions from Florida. Florida classifies drug offenses using broad ranges: 0-28g, 28-200g, 200-400g, and 400+g. The USSC data show a sharp 3.9 percentage point increase in cases with 200-400g convicted in a Florida federal district after 2010 (see Table 4, column 7 and Figure A8a). If the bunching in federal cases is due to state and local authorities sending more 280g cases to federal prosecutors, then there should be a mirrored decrease in the fraction of state-level cases in Florida with 200-400g.

I do not find a decrease in state convictions for 200-400g in general or by race. Figure 3a plots the share of all cocaine cases in Florida that are for offenses with 200-400g of cocaine by race. Columns 1 and 2 of Table 4 confirm this. The probability a state-level drug conviction is in the 200-400g range in Florida does not meaningfully change after 2010. This implies that shifting from state and local courts to federal courts cannot explain the sharp rise of cases at 280g in federal sentencing.[49]

Bunching by Law Enforcement Agency Sending Case to EOUSA. The EOUSA prosecutor case management files include a field that indicates the law enforcement agency that sends the case to the EOUSA. If the bunching at 280g is caused by a shift from state courts, then bunching should only be present in cases with state law enforcement involved. In Figure A8b, I plot the fraction of cases with 280-290g over time by the type of agency involved. I find that bunching at 280g is present in cases with state law enforcement involvement and in cases that are sent from Federal agencies. This is further evidence that the bunching at 280g after 2010 is not the result of state to federal case shifting.

3. Law Enforcement Discretion

Police can influence drug amounts by choosing the amount of drugs involved in "reverse sting" operations (in which agents sell drugs to offenders) or by extending traditional sting operations (in which agents buy drugs from offenders) until the total transacted amount is above the threshold (Honold 2014). I use data from local police and federal agents to explore the role of law enforcement discretion.

NIBRS, Local Law Enforcement Drug Seizures. Using a balanced panel of agencies in the NIBRS data on drug crime, I examine the distribution of drug seizure quantities. If local law enforcement is the source of bunching, I should observe an increase in bunching at 280-290g after 2010. Figure 3b plots the fraction of drug seizures with 280-290g over time and does not show an increase in drug seizures with 280-290g after 2010, in general or by race. These results are also shown in Columns 3 and 4 of Table 4.[50] In addition, only 5 incidents total are reported with 280-290g in the NIBRS after 2010. This suggests that discretion in local law enforcement and drug sting tactics cannot explain the bunching in drug amounts after 2010.

DEA STRIDE, Federal Law Enforcement Drug Seizures. I also test for bunching in drug quantities from the DEA's STRIDE database.[51] Figure 3c plots the share of cocaine exhibits with weights from 280-290g

[49] I find similar results using alternative sample restrictions in FL and using similar data from N. Carolina.

[50] This result is robust to using only states that have full coverage by 2012 (i.e. states in which all agencies are participating in NIBRS) and 90-100% coverage from at least 2008-2015 (US DOJ 2012).

[51] The analysis in this section uses unvalidated DEA data, and I claim authorship and responsibility for all inferences and

from 2000-2015. There is no increase in exhibits with 280-290g after 2010. Table 4 also shows this result. There are fewer than 20 total cocaine exhibits in the DEA data with 280-290g after 2010. This further suggests that local and federal law enforcement are not responsible for the bunching at 280g after 2010.

4. Prosecutorial Discretion

Bunching in Prosecutor Case Management Files. The EOUSA provides case-level data extracted from their internal case management system. Using this data, I test for bunching in the quantity of drugs recorded in that system. Figure 3d shows a sharp increase in the fraction of cases recorded with 280-290g after 2010.[52] Since I find no evidence of bunching in data from earlier stages, this suggests the bunching occurs once the case is in the hands of the prosecutor.

Table 4 indicates that the fraction of cases in 280-290g increases by 7.8 percentage points after 2010. This is twice the increase I find in the final sentencing data. This difference is likely driven by missing values in the EOUSA files. Re-coding each missing value as though it were not in the 280-290g range (i.e. equal to zero) yields an increase of about 2.4 percentage points after 2010, which is consistent with estimates from the sentencing data. The main results below are robust to missing value re-coding.[53,54]

The EOUSA data do not contain a field for race of the defendant. I impute race for cases that have a sentence month and year by using the racial composition of sentencing in each district-year-month from the USSC data. As before, I find an increase in 280-290g cases after 2010 and a larger increase in district-months with more black and Hispanic offenders. I also show that the disproportionate bunching by imputed race is robust to including prosecutor fixed effects (see Table A9).

Prosecutor-level Bunching Estimates. To further explore bunching by prosecutors, I use the ID of the lead attorney on each case and test for heterogeneity in bunching by attorney. Since each attorney only has a small number of cases and since I do not know the exact circumstances of each case, I cannot pinpoint "bad behavior" from any individual attorney. Also, cases are not randomly assigned to prosecutors at the federal level–assignment is based on many factors, such as caseloads or expertise. However, by estimating bunching for each attorney, I can calculate the fraction responsible for the observed bunching. Also, I can compare cases for bunching and non-bunching attorneys to further understand the source of the excess mass at 280g. Finally, I can provide summary statistics comparing bunching and non-bunching attorneys. Note, the EOUSA data do not include a field for prosecutor race.

---

conclusions that I draw from this information.

[52] See Figure A8c for a plot of bunching at 280g by the month the case is received.

[53] See Table A9 and Figure A8d-e for the main bunching results after re-coding the 280-290g binary variable as equal to zero when the drug weight is missing. For a more detailed discussion of missing values in the EOUSA data, see Online Appendix E of Tuttle (2022).

[54] Figure A8f examines bunching at 280g for cases received by the EOUSA before the FSA is signed into law. These cases are less likely to be influenced by offender or police responses to the FSA. For cases that are received by the EOUSA 60 days before the FSA but sentenced after the FSA, 2.7% are bunched at 280-290g. For cases that are received by the EOUSA 60 days before the FSA and sentenced before the FSA, 0.4% are bunched at 280-290g. The timing of bunching in these cases further suggests the increase in bunching at 280g is due to prosecutor decisions.

Prior to 2010, approximately 0.4% of cases were recorded as having 280-290g. For each attorney, I calculate the percentage of their cases with 280-290g after 2010. I classify an attorney as a "bunching attorney" if their bunching is greater than or equal to 0.4%.[55] For this analysis, I limit the sample to attorneys with 10 or more drug cases after 2010.[56] Most attorneys will be classified in the bunching category if they have at least one case in the 280-290g range after 2010.

The majority of these attorneys exhibit no bunching. Approximately 30.4% of prosecutors, however, do have a higher than normal percentage of cases with 280-290g after 2010. Drawing 50 samples (stratified on lead attorney ID and with replacement) from the data and re-calculating the fraction of bunching attorneys in each sample yields a standard error of 0.024 and a 90% confidence interval of about 26.4-34.3%.[57] Over 50% of bunching attorneys have two or more cases at 280-290g and over 25% have three or more cases at 280-290g. The fraction of bunching attorneys is also significantly different at the one percent level from the fraction calculated by randomly re-assigning cases to prosecutors (see Figure A9).

Since the EOUSA data do not contain a field for defendant race, I cannot compare racial disparities in charged amounts across bunching and non-bunching prosecutors. Using imputed race based on the USSC data, I find that bunching attorneys exhibit a greater degree of bunching in district-months with a higher share of black and Hispanic offenders. Removing bunching attorneys with 10+ cases from the sample reduces the size of the racial disparity estimated in Table A9 by 20%. Removing bunching attorneys with 5+ cases from the sample reduces it by 44% in total. This suggests that the attorneys I classify as bunching attorneys are important for overall racial disparities in bunching, although a substantial portion is due to attorneys with fewer than 5 cases after 2010. Finally, bunching attorneys charge higher drug amounts on average in district-months with more black and Hispanic offenders. For non-bunching attorneys, there is no relationship between district-month racial composition and average amount charged.[58] This supports the claim that the racial disparity in bunching is not driven by differential drug involvement by race.

The attorney-level bunching cannot be accounted for by district fixed effects. The within-district standard deviation in the 280-290g bunching metric is 0.1475, the between-district standard deviation is similar at 0.1472, and district fixed effects explain about 25% of the variance in the attorney-level bunching metric.

Further Evidence on Source of Excess Mass at 280g. In Table 6, I estimate the likelihood a case

[55] I find similar results using the district-level pre-2010 average to account for district fixed effects in cases at 280-290g. I can also use each attorney's pre-2010 behavior as their own benchmark to detect bunching post-2010. This also yields similar results but drastically limits the number of attorneys I can consider because few have a sufficient number of cases pre- and post-2010.

[56] Results are similar when using attorneys with 5 or more cases after 2010 or with 15 or more cases after 2010.

[57] While this statistic is only calculated for the 128 attorneys with 10 or more drug cases post-2010, this fraction holds for the entire data. In fact, bunching attorneys with 10+ cases post-2010 do not even account for half of the total bunching. Removing the bunching attorneys with 10+ cases post-2010 decreases the bunching estimate from 0.078 to 0.054. The majority of bunching at 280g is accounted for by prosecutors with fewer than 10 cases post-2010. Removing the bunching attorneys with 5+ cases post-2010 further decreases the bunching estimate from 0.078 to 0.039.

[58] Studying sentencing in the EOUSA data is difficult because the data on whether sentences are served consecutively or concurrently is missing in 80% of cases. Using USSC data, I find that districts with a greater racial disparity in bunching also have a greater overall racial disparity in crack-cocaine sentences.

is charged below 280g, with 280-290g, or above 290g for bunching versus non-bunching attorneys.[59] This echoes the approach that Goncalves and Mello (2021) use to estimate bunching in speeding tickets from police. For this analysis, I use two definitions of a bunching attorney: (1) attorneys who have an above-average share of cases with 280-290g post-2010 (using a leave-out mean) and (2) attorneys who have an above-average share of cases with 50-60g pre-2010.[60]

The key assumption here is that the non-bunching attorneys provide a counterfactual density since they are not responding to the mandatory minimum thresholds in the same way as the bunching attorneys. Comparing these two groups, I see that non-bunching attorneys (in both definitions) have more cases below 280g post-2010 than bunching attorneys and a similar number of cases above 290g post-2010. This provides further evidence, from different data and a different source of variation, that those attorneys who bunch at mandatory minimum thresholds are shading up the reported quantity of crack-cocaine.

Additional Evidence on Prosecutor-level Bunching. Next, I identify attorneys who switch from one federal district to another to test whether bunching is persistent across districts. Definition (2) from the above section is important for this analysis because there are few attorneys who switch districts and have a sufficient number of cases post-2010 in both districts. Table A11 shows that an attorney who bunches at the 10-year threshold in their first district is more likely to bunch at the 10-year threshold in their second district than an attorney who does not bunch at the 10-year threshold in their first district. Bunching at the 10-year threshold is a behavior that persists across districts, suggesting it is related to a characteristic of the prosecutor and not another actor in their district.

In Figure A10, I examine how other prosecutors in a district change their bunching behavior when a bunching prosecutor enters. Again relying mainly on definition (2), I find that when a bunching attorney switches into a new district, other attorneys in that district begin bunching more. Moreover, bunching in the district does not decrease once the bunching attorney leaves. This is highly suggestive evidence that the increase in bunching is not related to a temporary shift, such as competition among attorneys, but that it may be related to something more permanent, such as learning about techniques or developing new beliefs/norms.[61] The figure notes in the Online Appendix contain a more detailed discussion of the results.

In Table A12, I show that attorneys who bunch at 280-290g post-2010 also have more cases bunched at 28-29g (the 5-year mandatory minimum) post-2010 and more cases bunched at 50-60g pre-2010 (the pre-2010 10-year mandatory minimum). Likewise, attorneys who bunch at 50-60g pre-2010 also have more cases bunched at 28-29g post-2010 and 280-290g post-2010. One concern about the estimation of prosecutor-level bunching is that the variation across prosecutors could be due to noise alone, especially since I only require prosecutors to have 10 or more cases after 2010. These results that show prosecutor-level

[59] This approach ameliorates concerns about selection of cases changing after 2010 because it relies on post-2010 comparisons.

[60] I show that these results are robust to using attorneys with 15+ cases or 5+ cases. I also show that results are robust to using percent of cases bunched as the variable of interest. See Tables A10-A11

[61] Bunching at the 10-year mandatory minimum increased by 60% from 1988-90 to 2010, which is consistent with the practice of bunching being learned over time.

bunching is persistent across time, across districts, and across mandatory minimum thresholds provide strong evidence that the prosecutor-level bunching metric does contain a signal of prosecutor type.

Since cases are not randomly assigned to attorneys, one might expect that prosecutors with cases bunched at 280g are simply assigned more complex cases or cases with more drug involvement. I do not find that bunching attorneys are involved in more complex cases (see Table A12). In addition, bunching at 280g is weakly correlated with prosecutor tenure in the EOUSA.[62] Recall that bunching prosecutors are also more likely to have cases at the 28g threshold after 2010, and that bunching and non-bunching attorneys are equally likely to have cases above 290g. These results suggest that the bunching at 280g is not a result of bunching attorneys being assigned cases with more drug involvement or greater complexity.

C. The Impact of *Alleyne v. United States*

On January 14, 2013, the Supreme Court began hearing arguments in the case *Alleyne v. United States*. The petitioner, Allen Alleyne, argued that facts that increase the mandatory minimum sentence for a defendant are "elements" of the alleged crime and should be evaluated by a jury. In a 5-4 decision on June 17, 2013, the Court ruled in favor of Alleyne and issued a decision that changed the evidentiary standard for evidence related to mandatory minimum sentencing enhancements (Bala 2015).

Prior to this decision, evidence on drug quantities was presented to the judge during the "sentencing phase" of a trial. The judge would then decide, based on the legal standard of "a preponderance of evidence," whether the mandatory minimum sentence applied. The Supreme Court decision required that evidence that would raise the minimum sentence for a defendant be presented to the jury and evaluated based on the stricter legal standard of "beyond a reasonable doubt." Although most cases end in a plea deal, this decision changes the relative bargaining power of the prosecution and defense, particularly in cases with evidence that would not satisfy the stricter evidentiary standard.[63] I estimate how prosecutors reacted to this decision by comparing the change in bunching around June 17, 2013 to the change around June 17th in other years after 2010. If prosecutors are inflating drug amounts to levels that could not be supported at trial, then there will be a decrease in bunching for cases received after *Alleyne*.

Using the EOUSA data, I estimate the discontinuity in the prevalence of bunching for cases received around June 17, 2013 relative to cases received around June 17 in other years after 2010:

$$\begin{aligned}(\text{Recorded } 280-290g)_{it} = {} & \alpha_0 + \beta_1 \text{After}_{it} + \beta_2 \text{Days}_{it} + \beta_3 (\text{After}\times\text{Days})_{it} + \delta_1 (\text{After}\times\text{Year2013})_{it} \\ & + \delta_2 (\text{Days}\times\text{Year2013})_{it} + \delta_3 (\text{After}\times\text{Days}\times\text{Year2013})_{it} + D_{it} + \epsilon_{it}\end{aligned} \quad (6)$$

[62] Bunching attorneys do have cases with more defendants and witnesses, on average, but these are two mechanisms by which a prosecutor can increase the amount charged in the case. Notably, even in cases with only one defendant or cases with no witnesses, bunching attorneys are still more likely to have cases at 280g than non-bunching attorneys.

[63] For example, the defense can now threaten to challenge the amount at trial. A model of plea bargaining suggests prosecutors should only respond to credible threats, thus a decrease in bunching after *Alleyne* would suggest that the drug amount could be successfully challenged at trial in those cases. In other words, assume that juries find guilt if p(Guilty)>0.99 and a judge finds guilt if p(Guilty)>0.5. In that case, the threat is only credible if the p(Amount>=280g)>0.5 and <0.99.

where After_{it} is equal to one if case **i** is received after June 17th of year **t** but before January 1st of year **t+1** and equal to zero if case **i** is received before June 17th of year **t** but after January 1st of year **t**.[64] Days_{it} is the number of days from June 17th that case **i** is received, and Year2013_{it} is equal to one if case **i** is received in 2013 and equal to zero if it is received in 2011-2012 or 2014-2016. D_{it} represents day-of-week fixed effects. β_1 is the average discontinuity in the fraction of cases with 280-290g after June 17 from 2011-2016. Controlling for the discontinuity in other years is the standard and recommended approach in the regression discontinuity in time literature (Hausman and Rapson 2018). δ_1 is the discontinuity that is specific to June 17, 2013–the date of the *Alleyne* decision.[65]

Column 2 of Table 7 shows this result using a bandwidth of 130 days (the Imbens-Kalyanaraman optimal bandwidth for 2013) before and after June 17th in each year. The coefficient in the first row indicates that, on average, there is approximately no change in bunching after each June 17th from 2011-2016. The next coefficient shows the change in bunching that is specific to June 17, 2013. I find the fraction of cases recorded with 280-290g decreases discontinuously after the ruling in *Alleyne* on June 17, 2013.[66,67] Table A13 also shows that the decrease in bunching after *Alleyne* is robust to imputing missing values as zero and that missingness does not increase as a result of *Alleyne*. See Online Appendix E of Tuttle (2022) for further discussion of missing values in the EOUSA data. Figure A11 illustrates the large discontinuity in the fraction of cases with 280-290g around June 17, 2013, and shows robustness to additional bandwidth choices and choice of polynomial. Although it does not eliminate it entirely, it is clear that *Alleyne* at least somewhat reined in the practice of bunching.[68] This suggests that prosecutors were using discretion to build cases on

[64]This specification brings up a nuanced timing issue with the EOUSA and USSC data. In the EOUSA data, the drug quantity field does not necessarily display the final amount used for sentencing in the case. Instead, it is the amount entered into the case management system. Based on the user manual for the system, quantity is not a required field that must be updated throughout the life of the case. The user is first prompted to enter quantity when the case is opened. In practice, this is correlated with the final amount used in sentencing at the district-month level, and the patterns of bunching are similar in the EOUSA and USSC data. However, this difference is important for the analysis of *Alleyne*. In the EOUSA data, I examine how bunching differs based on when the case is received since that is when prosecutors are likely recording the amount in the system. In cases received and recorded before *Alleyne*, prosecutors often indicate that the quantity involved is at 280g. In cases received and recorded after, prosecutors are less likely to indicate that 280g were involved. *Alleyne*, however, will affect the final amount used at sentencing in cases received both before and after the decision as long as they are sentenced after *Alleyne*. For cases initiated before but sentenced after, prosecutors may need to file a superseding indictment or information that addresses drug quantity. This means that there should also be a discontinuity in bunching in final amount used at sentencing for cases sentenced before and after *Alleyne*. The USSC data does not include exact date of sentencing, but it does include sentence month and year. I use this data in Table A13 and show that bunching at final sentencing also falls, as we should expect, for cases sentenced shortly after *Alleyne*.

[65]The density of cases around this date is displayed in Figure A11a.

[66]The discontinuity for 2013 is at least twice as large as the next largest discontinuity from 1999-2016. See Figure A11b.

[67]I do not find a decrease in cases recorded with 280-290g after the announcement that the Court would hear the case (in October 2012) or after the oral arguments (in January 2013). The ruling in June 2013 was not clear from the outset. At the time, the NY Times referred to the case as a "murky area of sentencing law" on which the Court had issued "contradictory rulings." The announcement and the arguments alone would not provide sufficient evidence of whether the law would ultimately change.

[68]I do not test for a differential effect of *Alleyne* by race. The EOUSA data does not include defendant race and my imputation for race is at the monthly level. It is not clear *Alleyne* should reduce bunching at 280g differently by race. The idea is that bunched cases are based on weak evidence and that there are more of these cases for black and Hispanic defendants than white defendants. For those bunched cases, however, *Alleyne* should decrease their prevalence regardless of race. The effect is also complicated by whether juries are more or less biased than judges, since *Alleyne* increases the import of the jury. It is possible that *Alleyne* reduced racial sentencing disparities by decreasing the use of bunching at 280g. Using the post-2010 USSC data, I find the average sentence for black and Hispanic offenders fell from 7.97 to 7.85 years after *Alleyne* whereas the average sentence for white offenders rose

evidence that was unlikely to pass "beyond a reasonable doubt" scrutiny from juries.

In response to *Alleyne*, Attorney General Eric Holder released a memo in August 2013 instructing US attorneys to decline to charge quantities that trigger mandatory minimums in cases with non-violent defendants who have low criminal history. The decrease in bunching could be a result of this memo and not the Supreme Court decision. To address that concern, I narrow the bandwidth of the RD design to 60 days before/after June 17th. Even then, I find a discontinuous decrease in bunching (although the standard errors are larger). Also, using updated EOUSA data, I find that there is no change in bunching after May 12, 2017, the day Attorney General Jeff Sessions rescinded the August 2013 Holder memo. Finally, most cases bunched at 280-290g were not treated by the Holder memo. The Holder memo applied to defendants with criminal history scores of two or below, and over 70% of those bunched at 280-290g fall outside of that range. Didwania (2020) examines the effect of the Holder memo on sentencing, finding that sentences decreased by 6% among eligible offenders but racial disparities in sentencing remained static.

D. Discrimination and Alternative Explanations

Now, I consider four explanations for the racial disparity in bunching: taste-based racial discrimination, statistical discrimination based on race, racial differences in the cost (to the prosecutor) of bunching, and the possibility that the three previous explanations could be driven by a characteristic correlated with race rather than race itself. See Online Appendix D of Tuttle (2022) for a model that discusses these channels.

1. Other Offender Characteristics

First, I test the explanation that the racial disparity in bunching at 280g is driven by a characteristic correlated with race. I estimate equation (2) fully interacted with binary variables for several offender characteristics. This partially addresses concerns that white and black and Hispanic offenders are different on a wide range of other characteristics and that race may be a proxy for those characteristics. By estimating bunching by race and education, for example, I can compare black offenders with a college education to white offenders with a college education. If the racial disparity exists within education categories, this suggests that it is not driven by racial differences in education. In Table 8 and Figure 4a, I show that the racial disparity in bunching exists even within many observably similar groups.

To further explore the possibility that bunching is driven by an observable characteristic other than race, I do two additional tests. First, I examine the racial disparity in bunching within more narrow categories of criminal history score. In Figure 4b, I show that even among offenders with similar criminal history scores, there is a racial disparity in bunching. Second, I use the pre-2010 sentencing data and the available information about offenders and their offenses to predict a sentence for people with similar characteristics post-2010. In Figure 4c, I show that even among offenders within the same quantile of predicted sentence, there is a racial disparity in bunching. See Figures A12a-c for further analysis.

from 5.74 to 6.02 years. This corresponds to an 18% reduction in the racial sentencing gap over this time period.

The characteristics in the USSC data are only a subset of what the prosecutor observes. One concern is that black and Hispanic drug offenders may be more likely to operate in drug organizations or gangs.[69] Table A14 shows that bunching at 280g is slightly lower in states with high gang activity according to the DEA's 2009 Gang Threat Assessment, and that the racial disparity in bunching is the same in states with low versus high gang activity. Furthermore, the 2004 Survey of Inmates in Federal Correctional Facilities (SIFCF) indicates that black and Hispanic federal drug offenders are **less** likely to be a member of a drug organization prior to their arrest than white federal drug offenders (see Tables A15a-b). This suggests that racial differences in gang involvement would cause us to understate racial bias in bunching once focusing on defendants in federal court. Returning to the USSC data, although the amount charged is endogenous to the presence of a conspiracy charge, there is a racial disparity in bunching for offenders charged with conspiracy and for offenders not charged with conspiracy (see Table 8, column 8 and Figure A12d-e).[70] This suggests that differences in gang participation by race do not explain the racial disparity in bunching.

Although involvement in a criminal enterprise does not explain the disparity in bunching at 280g, there are, perhaps, other unobservable factors correlated with race that do. This paper cannot rule out every concern about unobservables. However, by leveraging the detail in the USSC data, the change in the mandatory minimum threshold after 2010, and the change in evidentiary standards in 2013, this paper rules out many alternatives. A compelling alternative explanation must not be captured by the detailed set of observables from the sentencing data, must contend with several pieces of evidence which suggest similar drug involvement by race, and must comport with the decline in bunching observed after evidentiary standards are increased specifically in mandatory minimum cases.

2. Costs to the Prosecutor of Bunching at 280g

In this section, I test the explanation that the racial disparity is due to racial differences in the costs to the prosecutor of bunching a case at 280g. First, I test whether a racial difference in defense counsel could explain the racial disparity in bunching.[71] The data do not include the offender's type of counsel in all years. From 1999-2002, black, Hispanic, and white crack-cocaine offenders are equally likely to be represented by private counsel.[72] Using data from the 1999-2002 USSC files, I construct each district's private counsel retention rate and tag districts as below or above median private counsel retention. I find that bunching and the racial disparity is similar in places with low and high private counsel retention (see Table A14).

Next, I consider whether the racial disparity in bunching can be attributed to judge bias. I am able to match approximately half of the cases in the EOUSA files to judge race and political party. For these cases, I

[69] Another possibility is that there are unobservable differences in violence, as perceived by the prosecutor. However, fewer than 0.1% of offenders in the sample are charged with a violent crime or flagged as violent in sentencing adjustments. Omitting this small sample of offenders does not change the main results. Likewise the presence of a weapons charge does not explain the racial disparity. Unfortunately, I do not observe location of the offense at a finer geography than federal district.

[70] Consistent with greater gang involvement in the SIFCF, white offenders are more likely to face a conspiracy charge overall.

[71] Agan, Freedman, and Owens (2019) demonstrate the importance of defense counsel type.

[72] 21.0% of white offenders, 22.7% of black offenders, and 21.7% of Hispanic offenders retain private counsel from '99-'02.

do not find any evidence that judge race or political party influences the probability a case is bunched at 280g (see Table A16). Also, unlike prosecutors, judges with a high share of cases at 280g post-2010 are not any more likely to have cases at 28g post-2010 or at 50g pre-2010 (see Table A17).[73] I also test whether district-level differences in costs of gathering evidence are related to the racial disparity in bunching at 280g. I find that the disparity is similar in districts with low and high fractions of cases declined due to "weak evidence" or "lack of resources" (see Table A14).[74,75]

3. Taste-based vs. Statistical Discrimination

Lastly, I consider taste-based and statistical discrimination. These explanations are difficult to disentangle and both are, ultimately, forms of racial discrimination. In addition, neither is benign in the eyes of the law. As Rehavi and Starr (2014) note, "otherwise-unconstitutional discrimination cannot be legally defended on the basis of statistical generalizations about group traits, regardless of their empirical support (*J.E.B. v. Alabama* ex rel T.B., 511 U.S. 127 1994)." However, given considerable interest in these models in economics, I discuss results which provide suggestive guidance about which form is at work in this setting.

A simple model of statistical discrimination would imply that prosecutors within the same district should be equally likely to bunch cases at 280g and that, after accounting for other offender characteristics, the racial disparity in bunching should decrease. I find that there is variation in the level of bunching across prosecutors within districts, and that the racial disparity exists within observably similar offender groups. While these results could be reconciled by a more detailed model of statistical discrimination, they suggest that standard statistical discrimination does not explain the racial disparity.

One potential explanation of these results is that some prosecutors have biased tastes against black and Hispanic drug offenders and believe they should be punished more harshly than white drug offenders. To explore the taste-based discrimination mechanism, I use a state-level measure of racial animus constructed by Stephens-Davidowitz (2014) based on intensity of Google searches including racial slurs in each state. I match this measure to the USSC data using the state of the federal district in which the offender is convicted. I take this measure of racial animus as a potentially valid measure of prosecutor tastes for several reasons: about half of government lawyers work in the same state they were born in (author's calculation from 2000 and 2010 publicly available Census samples), assistant US attorneys must reside in the district where they serve, and assistant US attorneys have a choice over where to apply.[76]

[73] Table A18 assuages concerns about juror bias because it shows that the disparity in bunching is not related to the share of black people in the district population (i.e. the potential jury pool). Also, bunching is highest before *Alleyne*, when jurors were less involved with the quantity determination.

[74] The EOUSA files only contain information about why a case is declined for about 60% of its cases.

[75] Another possibility is that if black and Hispanic defendants are less likely to plead guilty, prosecutors may use mandatory minimums to induce a plea. Table A19 shows the racial disparity in bunching exists in both districts where black and Hispanic defendants are more likely to plead guilty pre-2010 and in districts where they are less likely.

[76] Recall that *Alleyne v. US* made the jury more important in mandatory minimum cases after 2013. If juries are, on average, more racially biased than judges, then the effect of *Alleyne v. US* may be buffered by the increased racial bias of juries. I find that the fraction of cases at 280-290g in lower animus states fell by 40% from 2011-2012 to 2014-2017. In higher animus states, the fraction of cases at 280-290g fell by 20%. This suggests *Alleyne* was, in fact, less effective in states with higher racial animus.

Again, I estimate equation (2) fully interacted with a binary variable equal to one if the state where the offender is convicted is above the median on a measure of racial animus from Stephens-Davidowitz (2014) and equal to zero if it is below the median. If racism is correlated with state-level preferences for harsh sentencing, then I should find an effect for all offenders. However, I find that in states with a higher level of racial animus, bunching at 280g is more prevalent specifically for black and Hispanic offenders.[77] These results are in Tables 8-9.[78] Column 11 of Table 8 shows that in states with higher racial animus, black and Hispanic offenders are substantially more likely to be charged with an amount bunched at 280g.

Table 9 explores robustness. Columns 1-4 introduce individual and district-level controls interacted with post-2010 by race, and the relationship between animus and bunching remains.[79] Column 5 estimates the relationship between bunching and the continuous measure of state-level animus from Google Trends.[80] In column 6, I introduce a district-level measure of racial animus by aggregating implicit association test scores for people reporting an occupation of "lawyers, judges, and related workers."[81] Since many states contain multiple federal districts, I include state fixed effects interacted with post-2010 by race binary variables. The estimate is identified from within state variation in the IAT bias measure. I find the average IAT score of lawyers in a district is correlated with higher bunching for black and Hispanic offenders (p-value = 0.13).

VI. Conclusion

For federal drug crimes, a sharp increase in sentencing is triggered when the offense involves at or above a certain amount of drugs. In this paper, I show that there is substantial bunching at the point where the mandatory minimum sentence increases, and that bunching is disproportionately large for black and Hispanic offenders. I use the pre-2010 distribution of drug weights, when the threshold is at 50g instead of 280g, to show that the racial disparity in bunching at 280g post-2010 is conditional on drug involvement.

Since the bunching only appears in prosecutor case management data and at final sentencing but not in data on state-level convictions or drug seizures, it is likely a result of prosecutorial discretion. In fact, 20-30% of attorneys account for the bunching observed in the case management data. In addition, bunching becomes less prevalent among prosecutors following a Supreme Court decision that requires stricter evidentiary standards for drug quantity evidence. This, in addition to numerous other tests discussed above, suggests that prosecutors are shading drug amounts upward to induce longer sentences.

[77] Since the effect is only for black and Hispanic offenders, this implies that the disparity is not driven by black and Hispanic offenders being concentrated in states with higher animus. In states with higher animus, black and Hispanic offenders are bunched at higher rates. Over 70% of the overall disparity estimated in Table 2 remains after controlling for state-by-post fixed effects. Also, the main animus effect (e.g. an interaction between black and Hispanic, post-2010, and above median animus) is robust to the inclusion of state and state-by-post-2010 fixed effects. Table A18 analyzes bunching and district racial composition.

[78] The racial animus measure was developed to measure animus toward black people. I assume this is positively correlated with animus toward Hispanic people and focus on the pooled results. However, the estimates are similar if I exclude black offenders or Hispanic offenders. Excluding Hispanic offenders and using the full set of controls, the interaction between after 2010, black, and above median is 0.032 ($se = 0.011$) and the interaction between after 2010, white, and above median is -0.002 ($se = 0.023$).

[79] In Figures A12f-g, I apply an empirical Bayes shrinkage to state-level estimates of the disparity and show that is also correlated with racial animus (Chandra et al. 2016). Figure A12h plots bunching over time by state-level animus.

[80] The relationship is also present in the EOUSA data. Higher animus states have more "bunching" attorneys.

[81] The IAT for racial bias is designed to test implicit racial biases. See Project Implicit (Xu et al. 2019) for more detail.

Why do some prosecutors bunch black and Hispanic defendants at 280g more often than white defendants? The racial disparity cannot be explained by observable individual characteristics or district characteristics. Black and Hispanic crack-cocaine defendants are just as likely to retain private counsel as white defendants. Also, bunching at 280g is unrelated to judge race, political party, and the judge's share of cases at other mandatory minimum thresholds. Since only a subset of prosecutors practice bunching and there is variation across prosecutors within federal districts, a simple model of statistical discrimination does not apply either. This suggests the disparity may be the result of taste-based discrimination. In fact, I find the racial disparity in bunching at 280g is largest in states with higher levels of racial animus.

Finally, the bunching in drug weights and the racial disparity in bunching has meaningful implications for the racial sentencing gap. Despite affecting only 3.3% of cases post-2010, bunching at 280g can account for up to 6% of the racial disparity in crack-cocaine sentences. Although I focus on this specific empirical setting for identification benefits, I find that bunching disparities exist in other major drug types (Table A2), the prevalence of bunching in a district is correlated with appeals and reversals in that district (Table A8), and attorneys who bunch at 280-290g also bunch at other mandatory minimum ranges pre- and post-2010 (Table A12). This suggests that bunching at 280-290g is indicative of more widespread discretion in the legal system, particularly for drug cases. I find 3.4% of black and Hispanic crack-cocaine offenders are bunched at 280g after 2010 versus 1.3% of white crack-cocaine offenders. To give a conservative estimate, I assume that 3.4% and 1.3% of all drug cases from 1999-2015 were subject to similar discretion by race. That implies total costs of 1.3 billion dollars for black and Hispanic offenders versus 173 million dollars for white offenders. In terms of incarceration, the disparity implies 21,000 years sentenced due to this discretion for black and Hispanic offenders versus 2,900 years sentenced for white offenders.

References

Abrams, D., M. Bertrand, and S. Mullainathan. 2012. "Do Judges Vary in Their Treatment of Race?" *Journal of Legal Studies* 41 (2): 347-383.

Agan, A., J. Doleac, and A. Harvey. 2023. "Misdemeanor Prosecution." *Quarterly Journal of Economics*.

Agan, A., M. Freedman, and E. Owens. 2019. "Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense." *Review of Economics and Statistics* 1-51.

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

Alleyne v. United States, 570 U.S. 99. 2013. https://www.law.cornell.edu/supremecourt/text/11-9335.

Anbarci, N. and J. Lee. 2014. "Detecting Racial Bias in Speed Discounting: Evidence from Speeding Tickets in Boston." *International Review of Law and Economics* 38: 11–24.

Antonovics, K. L. and B. G. Knight. 2009. "A New Look at Racial Profiling: Evidence from the Boston Police Department." *Review of Economics and Statistics* 91 (1): 163–177.

Anwar, S. and H. Fang. 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 96 (1): 127–151.

Anwar, S., P. Bayer, and R. Hjalmarsson. 2012. "The Impact of Jury Race in Criminal Trials." *Quarterly Journal of Economics* 127 (2): 1017–1055.

Arnold, D., W. Dobbie, and C. S. Yang. 2018. "Racial Bias in Bail Decisions." *Quarterly Journal of Economics* 133 (4): 1885-1932.

Arora, A. 2018. "Too Tough on Crime? The Impact of Prosecutor Politics on Incarceration." Working paper.

Ba, B., D. Knox, J. Mummolo, R. Rivera. 2020. "The Impact of Racial and Ethnic Diversity in Policing." Working paper.

Bala, N. 2015. "Judicial Fact-Finding in the Wake of *Alleyne*." *New York University Review of Law & Social Change* 39 (1): 1–44.

Barkow, R. E. 2010. "Institutional Design & the Policing of Prosecutors: Lessons from Administrative Law." *Stanford Law Rev*iew 61 (4): 869.

Bibas, S. 2004. "Plea Bargaining Outside the Shadow of Trial." *Harvard Law Review* 2463–2547.

Bibel, D. 2015. "Considerations and Cautions Regarding NIBRS Data." *Justice Research and Policy* 16 (2): 185-194.

Bjerk, D. 2005. "Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion under Mandatory Minimum Sentencing." *Journal of Law and Economics* 48 (2): 591–625.
Bjerk, D. 2017a. "Mandatory Minimum Policy Reform and the Sentencing of Crack Cocaine Defendants: An Analysis of the Fair Sentencing Act." *Journal of Empirical Legal Studies* 14 (2): 370-396.
Bjerk, D. 2017b. "Mandatory Minimums and the Sentencing of Federal Drug Crimes." *Journal of Legal Studies* 46 (1): 93-128
BJS. 2019. "Federal Justice Statistics, 2015-2016: Full report." https://www.bjs.gov/content/pub/pdf/fjs1516.pdf.
Bohren, J. A., K. Haggag, A. Imas, and D. Pope. 2022. "Inaccurate Statistical Discrimination: An Identification Problem." Forthcoming at *Review of Economics and Statistics*.
BOP. 2020. "Prison Safety." https://www.bop.gov/about/statistics/statistics_prison_safety.jsp.
Boylan, R. T. 2005. "What Do Prosecutors Maximize? Evidence from Careers of U.S. Attorneys." *American Law and Economics Review* 7 (2): 379-402.
Chandra, A., A. Finkelstein, A. Sacarny, and C. Syverson. 2016. "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector." *American Economic Review* (106) 8: 2110-2144.
Close, B. R. and P. L. Mason. 2006. "After the Traffic Stops: Officer Characteristics and Enforcement Actions." *B.E. Journal of Economic Analysis and Policy* 6 (1).
Council of Europe. 1992. "Recommendation of the Committee of Ministers to Member States Concerning Consistency in Sentencing."
Cox, R. and J. Cunningham. 2020. "Financing the War on Drugs: The Impact of Law Enforcement Grants on Racial Disparities in Drug Arrests." Forthcoming at *Journal of Policy Analysis and Management*.
Davis, A. J. 2013. "In Search of Racial Justice: The Role of The Prosecutor." *NYU Journal of Legislation and Public Policy* 16 (4): 821-851.
Didwania, S. 2020. "Mandatory Minimums and Federal Sentences." Working paper.
Donohue, J. J., III. 2009. "Assessing the Relative Benefits of Incarceration: Overall Changes and the Benefits on the Margin." *Do Prisons make Us Safer? The Benefits and costs of the Prison Boom*, 269–342. New York: Russell Sage Foundation
Easterbrook, F. H. 1983. "Criminal Procedure as a Market System." *Journal of Legal Studies* 12 (2): 289–332.
EMCDDA (European Centre for Drugs and Drug Addiction). 2009. "The State of the Drug Problem in Europe."
Farrell, J. 2003. "Mandatory Minimum Firearm Penalties: A Source of Sentencing Disparity?" *Justice Research and Policy* 5 (1): 95-115.
Feigenberg, B. and C. Miller. 2020. "Racial Divisions and Criminal Justice: Evidence from Southern State Courts." Forthcoming at *American Economic Journal: Economic Policy*.
Fischman, J. B. and M. M. Schanzenbach. 2012. "Racial Disparities Under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums." *Journal of Empirical Legal Studies* 9 (4): 729-764.
Franklin, T. W. and T. K. S. Henry. 2020. "Racial Disparities in Federal Sentencing Outcomes: Clarifying the Role of Criminal History." *Crime & Delinquency* 66 (1): 3-32.
Glaeser, E. L., D. P. Kessler, and A. M. Piehl. 2000. "What Do Prosecutors Maximize? An Analysis of Federalization of Drug Crimes." *American Law and Economics Review* 2 (2): 259-290.
Goncalves, F. and S. Mello. 2021. "A Few Bad Apples? Racial Bias in Policing." *American Economic Review* 111 (5): 1406–1441.
Gordon, S. C. 2009. "Assessing Partisan Bias in Federal Public Corruption Prosecutions." *American Political Science Review* 103 (04): 534–54.
Grossman, G. M. and M. L. Katz. 1983. "Plea Bargaining and Social Welfare." *American Economic Review* 73 (4): 749-757.
Haney, C. 2001. "The Psychological Impact of Incarceration: Implications for Post-Prison Adjustment." Natl. Policy Conference.
Harrington, E. and H. Shafer. 2020. "Discerning Discretion: Estimating Prosecutor Effects at Criminal Sentencing." Working paper.
Hausman, C. and D. S. Rapson. 2018. "Regression Discontinuity in Time: Considerations for Empirical Applications." *Annual Review of Resource Economics* 10: 533–552.
Hoekstra, M. and C. Sloan. 2022. "Does Race Matter for Police Use of Force? Evidence from 911 Calls." *American Economic Review* 112 (3): 827–860.
Honold, D. 2014. "Note: Quantity, Role, and Culpability in the Federal Sentencing Guidelines." *Harvard Journal on Legislation* 51: 389-314.
Imbens, G., and K. Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933–959.
Jordan, A. 2021. "What Can Plea Bargaining Teach Us About Racial Bias in Criminal Justice?" Working paper.
Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." 2011. *Econometrica* 79 (3): 651–692.
Kleven, H. J.. 2016. "Bunching." *Annual Review of Economics* 8: 435–464.
Knorre, A. 2020. "Do Russian Police Fabricate Drug Offenses? Evidence From Seized Heroin's Weight Distribution." *Journal of Drug Issues* 50 (4): 378-391.
Knowles, J., N. Persico, and P. Todd. 2001. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109 (1): 203–229.
Knox, D., W. Lowe, and J. Mummolo. 2020. "Administrative Records Mask Racially Biased Policing." *American Political Science Review* 114(3): 619-637.
Krumholz, S. 2019. "The Effect of District Attorneys on Local Criminal Justice Outcomes." Working paper.
LaCasse, C. A. and A. Payne. 1999. "Federal Sentencing Guidelines and Mandatory Minimum Sentences: Do Defendants Bargain in the Shadow of the Judge?" *Journal of Law and Economics* 42 (S1): 245-269.
Landes, W. M. 1971. "An Economic Analysis of the Courts." *Journal of Law and Economics* 14 (1): 61-107.
Lee, D. S. and J. McCrary. 2017. "The Deterrence Effect of Prison: Dynamic Theory and Evidence." *Advances in Econometrics* 38.
Lepage, L-P. 2020. "Are criminals strategic? Offender responses to drug sentencing cutoffs." *Labour Economics* 66.
Light, M. T. 2022. "The Declining Significance of Race in Criminal Punishment: Evidence from US Federal Courts." *Social Forces* 100 (3): 1110–1141.
Luh, E. 2022. "Not So Black and White: Uncovering Racial Bias from Systematically Misreported Trooper Reports." Working paper.

Lynch, M. *Hard Bargains: The Coercive Power of Drug Laws in Federal Court.* New York: Russell Sage Foundation, 2016.
Mueller-Smith, M. 2015. "The Criminal and Labor Market Impacts of Incarceration." Working paper.
Mueller-Smith, M. and K. Schnepel 2021. "Diversion in the Criminal Justice System ." *Review of Economic Studies* 88 (2): 883–936.
Mustard, D. B. 2001. "Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts." *Journal of Law and Economics* 44: 285–314.
Nagrecha, M. 2021. "Disparities and Discrimination in the European Unionâs Criminal Legal Systems."
Nyhan, B. and M. M. Rehavi. 2017. "Tipping the Scales? Testing for Political Influence on Public Corruption Prosecutions." Working paper.
Ouss, A. and M. Stevenson. 2022. "Does Cash Bail Deter Misconduct?" Forthcoming at *American Economic Journal: Applied Economics*.
Park, K. 2017. "Do Judges Have Tastes for Discrimination? Evidence From Criminal Courts." *Review of Economics and Statistics* 99 (5) 810-823.
Pfaff, J. *Locked In: The True Causes of Mass Incarceration and How to Achieve Real Reform*. New York: Basic Books, 2017.
PPI. 2019. "Mass Incarceration: The Whole Pie 2019." https://www.prisonpolicy.org/reports/pie2019.html.
Rehavi, M. M., and S. B. Starr. 2012. "Racial Disparity in Federal Criminal Sentences and Its Sentencing Consequences." Univ. of Michigan Law School Program in Law and Economics Working Paper Series. Working paper no. 12-002.
Rehavi, M. M., and S. B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6):1320–54.
Reinganum, J. F. 1988. "Plea Bargaining and Prosecutorial Discretion." *American Economic Review* 78 (4): 713-728.
Rhodes, W., R. Kling, J. Luallen, and C. Dyous. 2015. "Federal Sentencing Disparity: 2005-2012." BJS Working Paper Series.
Robert Wood Johnson Foundation (RWJF). 2018. "Discrimination in America: Final Summary." https://www.rwjf.org/content/dam/farm/reports/surveys_and_polls/2018/rwjf443620.
Shermer, L. O. and B. D. Johnson. 2010. "Criminal Prosecutions: Examining Prosecutorial Discretion and Charge Reductions in U.S. Federal District Courts." *Justice Quarterly* 27 (3): 1-37.
Silveira, B. S. 2017. "Bargaining With Asymmetric Information: An Empirical Study of Plea Negotiations." *Econometrica* 85 (2): 419-452.
Simon, J. 2007. *Governing Through Crime: How the War on Crime & American Democracy Created a Culture of Fear*. Oxford Univ. Press
Sklanskz, D. A. 2018. "The Problems with Prosecutors." *Annual Review of Criminology* 1: 451–469.
Sloan, C. 2022. "Do Prosecutor and Defendant Race Pairings Matter? Evidence from Random Assignment." Working paper.
Small, M. L. and D. Pager. 2020. "Sociological Perspectives on Racial Discrimination." *Journal of Economic Perspectives* 34 (2): 49-67.
Spohn, C., J. Gruhl, and S. Welch. 1987. "The Impact of the Ethnicity and Gender of Defendants on the Decision to Reject or Dismiss Felony Charges." *Criminology* 25 (1): 175–192.
Stephens-Davidowitz, S. 2014. "The Cost of Racial Animus on a Black Candidate: Evidence using Google Search Data." *Journal of Public Economics* 118: 26-40.
Stuntz, W. J. 2004. "Plea Bargaining and Criminal Law's Disappearing Shadow." *Harvard Law Review* 117 (8): 2548.
Taifa, N.. 1998. "An Insider's Perspective on the Crack Cocaine Controversy." *Federal Sentencing Reporter* 10 (4): 200-203.
The Hamilton Project. 2016. "Twelve Facts about Incarceration and Prisoner Reentry." Economic Facts, The Hamilton Project, The Brookings Institution, Washington, DC
Tonry, M. 2012. "Prosecutors and Politics in Comparative Perspective." *Crime and Justice* 41 (1): 1-33.
Traxler, C., F. G. Westermaier, and A. Wohlschlegel. 2018. "Bunching on the Autobahn? Speeding Responses to a 'Notched' Penalty Scheme." *Journal of Public Economics* 157: 78–94.
Travova, E. 2020. "Under Pressure? Performance Evaluation of Police Officers as an Incentive to Cheat: Evidence from Drug Crimes in Russia." Working paper.
Tuttle, C. 2022. "Online Appendix for 'Racial Disparities in Federal Sentencing: Evidence from Drug Mandatory Minimums' November 2022 manuscript." https://codytuttle.github.io/tuttle_mandatory_minimums_online_appendix_2022.pdf
Ulmer, J. T., M. C. Kurlychek, and J. H. Kramer. 2007. "Prosecutorial Discretion and the Imposition of Mandatory Minimum Sentences." *Journal of Research in Crime and Delinquency* 44 (4): 427-458.
Ulmer, J., N. Painter-Davis, and L. Tinik. 2014. "Disproportional Imprisonment of Black and Hispanic Males: Sentencing Discretion, Processing Outcomes, and Policy Structures." *Justice Quarterly* 33 (4): 642-681.
US DOJ. 2011. "Review of The Department of Justice's International Prisoner Transfer Program." https://oig.justice.gov/reports/2011/e1202.pdf.
US DOJ. 2012. "NIBRS Participation by State." https://ucr.fbi.gov/nibrs/2012/resources/nibrs-participation-by-state.
USSC. 2015a. "Report to the Congress: Impact of the Fair Sentencing Act of 2010." https://www.ussc.gov/research/congressional-reports/2015-report-congress-impact-fair-sentencing-act-2010.
USSC. 2015b. "United States Sentencing Commission (USSC) Guidelines Manual." https://www.ussc.gov/guidelines/guidelines-archive/2015-guidelines-manual.
USSC. 2017. "Demographic Differences in Sentencing: An Update to the 2012 Booker Report." https://www.ussc.gov/research/research-reports/demographic-differences-sentencing.
West, J. 2018. "Racial Bias in Police Investigations." Working paper.
Wright, R. 2006. "Federal or State? Sorting as a Sentencing Choice." *Criminal Justice* 21 (2): 16-21.
Xu, F., B. Nosek, A. Greenwald, K. Ratliff, Y. Bar-Anan, E. Umanksy, M. Banaji, N. Lofaro, C. Smith, J. Axt. 2019. "Project Implicit Demo Website Datasets."
Yang, C. S. 2015. "Free at Last? Judicial Discretion and Racial Disparities in Federal Sentencing." *Journal of Legal Studies* 44 (1): 75-111.
Yang, C. S. 2016. "Resource Constraints and the Criminal Justice System: Evidence from Judicial Vacancies." *American Economic Journal: Economic Policy* 8 (4): 289–332.

Tables and Figures

Table 1. Summary Statistics for USSC Sentencing Data.

	1999-2010	2011-2015
Black or Hispanic	0.921	0.940
	(0.270)	(0.238)
Age (in years)	31.237	34.148
	(8.493)	(8.729)
Male	0.917	0.917
	(0.276)	(0.276)
College or more	0.127	0.147
	(0.333)	(0.354)
High school or more	0.512	0.599
	(0.500)	(0.490)
Not US citizen	0.043	0.031
	(0.203)	(0.173)
Number of dependents	1.840	1.981
	(1.910)	(2.030)
Weapon involved	0.260	0.292
	(0.439)	(0.455)
Number of total current offenses	2.364	2.381
	(1.161)	(1.271)
Criminal history points	5.772	6.535
	(5.487)	(5.593)
Drug weight (in grams)	100.898	115.448
	(155.214)	(168.581)
Sentence (in years)	9.226	7.807
	(6.599)	(5.837)
Observations	41,253	9,020

Notes. The table above describes defendants found in the USSC sentencing data pre- and post-2010. The mean value of each variable is reported with standard deviations in parentheses. The statistics above are derived from the cleaned USSC data in which the following cases are removed: cases with missing drug weight values (including those cases with weights coded as a range), cases with reported problems in the drug weight or sentencing variables, cases where judges change or do not accept the findings of fact for drug weights, and cases at and above 1000g. Data construction is discussed in greater detail in Online Appendix E of Tuttle (2022). See Tables A1a-A1b for summary statistics for NIBRS, DEA, FL Inmate, and EOUSA data.

Table 2. Effect of Changing Mandatory Minimum Threshold on Bunching at 280-290g.

	Pr(280-290g Crack-Cocaine Recorded)		
	(1)	(2)	(3)
After 2010	0.0332***		0.0674***
	(0.00204)		(0.0129)
After 2010 x White		0.0135**	
		(0.0056)	
After 2010 x Black or Hispanic		0.0344***	
		(0.0021)	
Constant	0.00487***	0.0031***	0.00330***
	(0.000343)	(0.0010)	(0.00125)
P-value: W (White) = BH (Black or Hispanic)	-	0.0004	-
Trial Cases Only	No	No	Yes
Observations	50,273	50,273	2,515

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes on data construction. The row "P-value: W (White) = BH (Black or Hispanic)" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." In the remaining tables, I abbreviate the label to "P-value: W=BH." Specifications with the race and after 2010 interactions also include a binary variable equal to one for black and Hispanic offenders and equal to zero for white offenders. Black and Hispanic offenders are slightly more likely (0.2 percentage points) to be charged in the 280-290g range prior to 2010. This approach addresses the small pre-existing difference by estimating the increase in probability after 2010 by race. Coefficients are estimated from the following regression for Column 1:

$$(1) \quad (\text{Charged } 280-290g)_{it} = \alpha_0 + \beta_1 \text{After2010}_{it} + \epsilon_{it}$$

and the following regression for Column 2:

$$(2) \quad (\text{Charged } 280-290g)_{it} = \alpha_0 + \beta_1(\text{After2010} \times \text{White})_{it} + \beta_2(\text{After2010} \times \text{BlackOrHispanic})_{it} + \text{BlackOrHispanic}_{it} + \epsilon_{it}$$

Column 3 re-estimates equation (1) excluding cases that end in a plea deal (i.e. trial cases only). I do not re-estimate equation (2) on the trial-only sample because there are zero white offenders with 280-290g in trial cases after 2010. Results are robust to removing cases in DC, which has a fundamentally different prosecution system that relies more heavily on federal court. In this table and further analyses, I estimate both β_1 and β_2 to show the change in bunching by race post-2010 as transparently as possible. One could instead estimate the interaction After2010 x BlackOrHispanic to show the change for black or Hispanic offenders relative to white offenders. If I do this, the coefficient is equal to $\beta_2 - \beta_1$ and is statistically significant, as indicated by the row in this table labeled "p-value." See Tables A3-A6 for additional robustness tests.
*** p<0.01, ** p<0.05, * p<0.1

Table 3. "Missing Mass" in the Distribution of Drug Amounts, Comparing Pre- and Post-2010 Distributions

Panel A. Analysis of Changes in the 0-100g Range.					
	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	0.0199***	-0.0745***	0.0347***	-0.0052*	-0.0084**
	(0.0039)	(0.0049)	(0.0040)	(0.0029)	(0.0037)
Constant	0.1130***	0.2967***	0.1104***	0.0708***	0.1244***
	(0.0016)	(0.0022)	(0.0015)	(0.0013)	(0.0016)
Predicted Change from Conceptual Model	Increase	Ambiguous	Increase	Decrease	Decrease
Observations	50,273	50,273	50,273	50,273	50,273
Panel B. Analysis of Changes in the 100-1000g Range.					
	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
	(6)	(7)	(8)	(9)	(10)
After 2010 (Actual Change)	-0.0131***	0.0332***	0.0043*	0.0025	0.0067***
	(0.0045)	(0.0020)	(0.0025)	(0.0017)	(0.0020)
Constant	0.1903***	0.0049***	0.0437***	0.0206***	0.0254***
	(0.0019)	(0.0003)	(0.0010)	(0.0007)	(0.0008)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	50,273	50,273	50,273	50,273	50,273

Notes. Robust standard errors estimated jointly by seemingly unrelated regression in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes on data construction. The predicted change from the conceptual model of prosecutor behavior in Section II.B is displayed in the row labeled "predicted change from conceptual model." Coefficients are estimated from the following regression for each range:

$$(3) \quad (\text{Charged X-Yg})_{it} = \alpha_0 + \beta_1 \text{After2010}_{it} + \epsilon_{it}$$

Results are robust to removing cases in DC (see Table 2 notes). Table A7 displays versions of missing mass results with time trend interactions and interactions by race.
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Bunching Analysis for Potential Mechanisms

Panel A. Analysis of Bunching in State Convictions and in Drug Seizures					
	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)	Pr(280-290g)	Pr(280-290g)
	(1)	(2)	(3)	(4)	(5)
After 2010	0.00005		-0.0002***		-0.0006***
	(0.0005)		(.0001)		(0.0002)
After 2010 x White		0.0004		-0.0001	
		(0.0011)		(0.0001)	
After 2010 x Black or Hispanic		0.0002		-0.0003***	
		(0.0005)		(0.0001)	
Constant	0.0051***	0.0085***	0.0004***	0.0002***	0.0010***
	(0.0003)	(0.0005)	(0.00005)	(0.0001)	(0.0001)
Data Analyzed	FL Convictions	FL Convictions	Drug Seizures, NIBRS	Drug Seizures, NIBRS	Drug Evidence, DEA STRIDE
Drugs Included	Cocaine, all types	Cocaine, all types	Crack-cocaine	Crack-cocaine	Cocaine, all types
P-value: W = BH	-	0.8148	-	0.2383	-
Observations	214,573	214,573	191,677	191,677	100,306
Panel B. Analysis of Bunching in Prosecutor Case Files and Final Sentencing					
	Pr(280-290g)	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)	Pr(280-290g)
	(6)	(7)	(8)	(9)	(10)
After 2010	0.0783***	0.0389***		0.0332***	
	(0.00561)	(0.0128)		(0.00204)	
After 2010 x White			-0.0050		0.0135**
			(0.0288)		(0.0056)
After 2010 x Black or Hispanic			0.0434***		0.0344***
			(0.0132)		(0.0021)
Constant	0.0039***	0.1031***	0.1186***	0.00487***	0.0031***
	(0.0004)	(0.00745)	(0.0158)	(0.000343)	(0.0010)
Data Analyzed	EOUSA Case Management System	USSC Sentencing, FL only	USSC Sentencing, FL only	USSC Sentencing	USSC Sentencing
Drugs Included	Crack-cocaine	Cocaine, all types	Cocaine, all types	Crack-cocaine	Crack-cocaine
P-value: W = BH	-	-	0.0941	-	0.0004
Observations	19,363	6,605	6,605	50,273	50,273

Notes. Robust standard errors in parentheses. When possible, the specifications above use a sample of offenses with drug amounts between 0 grams and 1000 grams. Analyses of state-level drug convictions do not make this restriction since the state reports broad drug weight categories instead of specific amounts. When broad categories (e.g. 200-400g) are analyzed, a linear trend in year is included. The row "P-value: W= BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." In Panel A: columns 1-2 show an analysis of reported drug amounts for state-level drug convictions in Florida, columns 3-4 show an analysis of weights for seized drugs reported to the FBI through the National Incident Based Reporting System, and column 5 shows an analysis of weights for drugs sent to DEA laboratories. In Panel B: column 6 shows an analysis of weights recorded in case management files from the Executive Office of the US Attorney, columns 7-8 show an analysis of weights from USSC sentencing data for federal convictions in FL using broad drug categories and all types of cocaine, and columns 9-10 show the main bunching results from Table 2 for all federal crack-cocaine convictions in the USSC sentencing data. Coefficients in columns 1, 3, 5, 6-7, and 9 are estimated from the regression in equation (1) of Table 2, with a linear time trend included for columns 1 and 7 (the broad drug categories). Coefficients in columns 2, 4, 8, and 10 are estimated from the regression in equation (2) of Table 2, with a linear time trend included for columns 2 and 8. Results are robust to removing cases in DC (see Table 2 notes).
*** p<0.01, ** p<0.05, * p<0.1

Table 5. Offender Behavior by Race

Panel A. Drug Seizures Before and After the Fair Sentencing Act in 2010

	Weight	Pr(280-290g)	Weight	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-280g)	Pr(270-280g)	Pr(280-290g)	Pr(>290g)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After 2010 x White			0.0768	0.0342***	-0.0298***	0.0000	-0.0058***	-0.0000	-0.0000	0.0015**
			(0.6040)	(0.0041)	(0.0037)	(0.0017)	(0.0012)	(0.0000)	(0.0001)	(0.0007)
After 2010 x Black			-2.9470***	0.0531***	-0.0264***	-0.0077***	-0.0171***	-0.0001***	-0.0002**	-0.0016***
			(0.2774)	(0.0029)	(0.0026)	(0.0011)	(0.0010)	(0.0001)	(0.0001)	(0.0004)
Black	1.716***	0.0001	2.4062***	-0.0951***	0.0707***	0.0101***	0.0131***	0.0001***	0.0001	0.0009**
	(0.265)	(0.0001)	(0.2867)	(0.0026)	(0.0024)	(0.0010)	(0.0009)	(0.0001)	(0.0001)	(0.0004)
Constant	10.266***	0.0003**	9.8706***	0.7280***	0.2031***	0.0345***	0.0303***	0.0001	0.0003**	0.0038***
	(0.436)	(0.0001)	(0.4458)	(0.0041)	(0.0037)	(0.0016)	(0.0015)	(0.0001)	(0.0001)	(0.0006)
Observations	191,677	191,677	191,677	191,677	191,677	191,677	191,677	191,677	191,677	191,677
P-value: W = B	-	-	0.0000	0.0002	0.4433	0.0001	0.0000	0.0282	0.2444	0.0002

Panel B. Drug Use and Drug Selling After the Fair Sentencing Act in 2010

	Ever Use Crack	Sold Drugs in Past Year	Use Crack & Sold Drugs
	(11)	(12)	(13)
After 2010 x White	0.0019**	-0.0009**	-0.0007***
	(0.0009)	(0.0005)	(0.0002)
After 2010 x Black or Hispanic	-0.0053***	-0.0031***	-0.0010***
	(0.0015)	(0.0009)	(0.0003)
Black or Hispanic	0.0033***	0.0039***	-0.0009***
	(0.0012)	(0.0007)	(0.0003)
Constant	0.0342***	0.0145***	0.0037***
	(0.0005)	(0.0007)	(0.0001)
Observations	763,335	762,322	762,054
P-value: W = BH	0.0000	0.0257	0.3350

Notes. **Panel A:** Robust standard errors estimated jointly by seemingly unrelated regression in parentheses. This analysis uses the weights of seized drugs reported to the FBI through the National Incident Based Reporting System. Ethnicity is not consistently recorded in NIBRS over this time period. As such, I refer to offenders as black or white, omitting the Hispanic label used in previous analyses. Columns 1-2 show the relationship between race of offender and drug weight seized, in general. Column 3 shows how the weight of an offender's seized drugs changes by race after 2010. Columns 4-10 show how the probability an offender's seized drugs are in a certain bin changes by race after 2010. All specifications include state fixed effects and controls for age and sex. The row "P-value: W=B" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black." Coefficients in column 1 are estimated from the following regression: (5) $\text{Weight}_i = \alpha_0 + \beta_1 \text{Black}_i + X_i + Z_s + \epsilon_i$. The coefficients in column 2 are estimated from the same specification with a binary variable for the 280-290g range as the dependent variable. Coefficients in column 3 are estimated from the following regression: (6) $\text{Weight}_{it} = \alpha_0 + \beta_1(\text{Black} \times \text{After2010})_{it} + \beta_2(\text{White} \times \text{After2010})_{it} + \text{Black}_i + X_i + Z_s + \epsilon_{it}$ The coefficients in columns 4-10 are estimated from the same specification with binary variables for the range of interest as the dependent variable. **Panel B:** Robust standard errors in parentheses. This analysis uses data from the National Survey on Drug Use and Health. Column 11 shows that the fraction of respondents answering "yes" to the question, "have you ever, even once, used crack-cocaine?" does not increase after 2010. Column 12 shows that the fraction of respondents answering a number greater than zero to the question, "how many times have you sold illegal drugs in the past 12 months?" does not increase after 2010. Column 13 shows that the fraction of people answering yes to both of these questions does not increase after 2010. All specifications use year-specific sampling weights. The row "P-value: W=BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." Coefficients are estimated from the following regression: (7) $\text{Outcome}_{it} = \alpha_0 + \beta_1(\text{BlackOrHispanic} \times \text{After2010})_{it} + \beta_2(\text{White} \times \text{After2010})_{it} + \text{BlackOrHispanic}_i + \epsilon_{it}$.
*** p<0.01, ** p<0.05, * p<0.1

Table 6. Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors

Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)
Atty. Bunches at 280-290g Post-2010	-0.1137*	0.1491***	-0.0354
	(0.0659)	(0.0435)	(0.0463)
Constant	0.8912***	0.0272***	0.0816*
	(0.0432)	(0.0076)	(0.0436)
Observations	971	971	971
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Atty. Bunches at 50-60g Pre-2010	-0.0785***	0.0575***	0.0211
	(0.0254)	(0.0172)	(0.0168)
Constant	0.9359***	0.0233**	0.0408***
	(0.0170)	(0.0105)	(0.0133)
Observations	1,135	1,135	1,135

Notes. Standard errors clustered at the prosecutor level and estimated jointly by seemingly unrelated regression in parentheses. Bootstrapping standard errors to account for "atty. bunches" being an estimated regressor yields similar results. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(8) \quad (\text{Charged X-Yg})_i = \alpha_0 + \beta_1 \text{AttyBunchesAt280}g_i + \epsilon_i$$

where AttyBunchesAt280g is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g post-2010 is at or below the average fraction of 280-290g cases pre-2010). These regressions are restricted to post-2010 cases and to prosecutors with 10+ cases post-2010. This is a necessarily arbitrary restriction that is limited by the fact that most prosecutors do not handle many crack-cocaine cases. See Table A10 for robustness to adjusting this 10+ case restriction. Note, to avoid a mechanical relationship in column (2), I use leave-out-means to classify bunching attorneys. Coefficients in panel B are estimated from the following regression for each range:

$$(9) \quad (\text{Charged X-Yg})_i = \alpha_0 + \beta_1 \text{AttyBunchesAt50}g_i + \epsilon_i$$

where AttyBunchesAt50g is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 50g definition (i.e. the fraction of their cases that are from 50-60g pre-2010 is above the average fraction of 50-60g cases post-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 50-60g is at or below the average fraction of 50-60g cases post-2010). These regressions are restricted to post-2010 cases and to prosecutors with 10+ cases pre-2010. Results are robust to removing cases in DC (see Table 2 notes). See Tables A9-10 for additional robustness tests.

*** p<0.01, ** p<0.05, * p<0.1

Table 7. Change in Bunching by Prosecutors after *Alleyne v. United States* Decision

	Pr(Case Recorded with 280-290g)			
	(1)	(2)	(3)	(4)
After June 17th, 2011-2016	0.0070	-0.0049	0.0041	-0.0206
	(0.0260)	(0.0284)	(0.0295)	(0.0406)
After June 17th, 2013	-0.1735**	-0.1518*	-0.1433	-0.1289
	(0.0813)	(0.0920)	(0.0935)	(0.1246)
Constant	0.1624	0.1626	0.1576	0.2093
	(0.1522)	(0.1519)	(0.1520)	(0.1776)
Bandwidth	±150 days	±130 days	±120 days	±60 days
Observations	1,937	1,672	1,513	754

Notes. Standard errors clustered at the date the case is received in parentheses. The estimates in this table are based on the EOUSA data. The coefficients above are estimated from the following regression discontinuity style model:

$$\begin{aligned}(10)\quad (\text{Recorded } 280-290g)_{it} = {} & \alpha_0 + \beta_1 \text{AfterJune17}_{it} + \beta_2 \text{DaysFrom}_{it} + \beta_3 (\text{AfterJune17}\times\text{DaysFrom})_{it} \\ & + \delta_1 (\text{AfterJune17}\times\text{Year2013})_{it} + \delta_2 (\text{DaysFrom}\times\text{Year2013})_{it} \\ & + \delta_3 (\text{AfterJune17}\times\text{DaysFrom}\times\text{Year2013})_{it} + D_{it} + \epsilon_{it}\end{aligned}$$

where AfterJune17 is a binary variable equal to one for cases received after June 17th in each year, DaysFrom, the running variable, is the date the case was received centered at zero on June 17th, and Year2013 is equal to one for cases received in 2013 (the year *Alleyne* is decided). In addition, all specifications above include day-of-week fixed effects, D_{it}, for the day the case is received. The ±130 day bandwidth is selected from the Imbens-Kalyanaraman optimal bandwidth procedure for the year 2013. Figure A11 shows graphical evidence of the discontinuity in bunching around June 17, 2013. Results are robust to removing cases in DC (see Table 2 notes). See Table A13 for additional robustness tests.
*** p<0.01, ** p<0.05, * p<0.1

Table 8. Degree of Bunching Post-2010 by Race and Offender Characteristics.

	Pr(280-290g)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
After '10 x White (W)	0.0177**	0.0070	0.0164	0.0068	0.0137**	0.0143**	0.0163**	0.0109*	-0.0014*	0.0163**	0.0090
	(0.0070)	(0.0066)	(0.0104)	(0.0059)	(0.0057)	(0.0065)	(0.0080)	(0.0061)	(0.0008)	(0.0079)	(0.0092)
After '10 x Black or Hispanic (BH)	0.0349***	0.0245***	0.0389***	0.0313***	0.0338***	0.0301***	0.0415***	0.0307***	0.0089***	0.0427***	0.0197***
	(0.0023)	(0.0073)	(0.0039)	(0.0039)	(0.0022)	(0.0024)	(0.0039)	(0.0025)	(0.0015)	(0.0150)	(0.0072)
After '10 x W x Char.	-0.0196***	0.0111	-0.0042	0.0122	-0.0137**	-0.0039	-0.0062	0.0100	0.0300**	-0.0082	0.0070
	(0.0073)	(0.0105)	(0.0122)	(0.0106)	(0.0057)	(0.0127)	(0.0111)	(0.0140)	(0.0117)	(0.0110)	(0.0122)
After '10 x BH x Char.	-0.0040	0.0105	-0.0068	0.0042	0.0197	0.0145***	-0.0109**	0.0111**	0.0648***	-0.0133	0.0191*
	(0.0061)	(0.0076)	(0.0047)	(0.0047)	(0.0148)	(0.0050)	(0.0047)	(0.0047)	(0.0052)	(0.0160)	(0.0114)
Constant	0.0033***	0.0021	0.0043**	0.0015	0.0031***	0.0027***	0.0017*	0.0038***	0.0014*	0.0034**	0.0058**
	(0.0011)	(0.0015)	(0.0017)	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0013)	(0.0008)	(0.0016)	(0.0022)
Characteristic	College	Male	Above Med. Age	Has Dependents	Not a Citizen	Weapon	Above Med. Crim. Hist. Points	Above Med. # of Counts	Conspiracy Charge	State Above Med. % of Black and Hispanic Cases	State Above Med. Racial Animus
P-value: W = BH	0.0195	0.0739	0.0415	0.0005	0.0224	0.0085	0.0048	0.0027	0.0000	0.1204	0.3799
P-value: W+Char. = BH+Char.	0.0000	0.0449	0.0038	0.0717	0.0036	0.0007	0.0116	0.1161	0.0004	0.0113	0.0381
Observations	50,273	50,273	50,273	50,273	50,273	50,273	50,273	50,273	50,273	50,273	50,090

Notes. Robust standard errors in parentheses for columns 1-9. Standard errors clustered at the state level are in parentheses for columns 10 and 11. "Characteristic" or "Char." represents a binary variable that is an offender or case characteristic. The estimates in this table are based on the USSC data. The specific offender characteristic of interest is noted in the "Characteristic" row. For example, when the "Characteristic" is "College", then "Characteristic" is equal to one if the offender's educational attainment is college or more and is equal to zero if the offender's educational attainment is less than college. See Table 1 for notes on data construction. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "P-value: W+Char. = BH+Char." reports the p-value from a test of the null hypothesis that the combined coefficients on "(After 2010 x White)+(After 2010 x White x Characteristic)" is equal to the combined coefficients on "(After 2010 x Black or Hispanic)+(After 2010 x Black or Hispanic x Characteristic)." Male is equal to one if the offender is male and equal to zero if not. Above median age is equal to one if the offender is above the median age for offenders and equal to zero if not. Has dependents is equal to one if the offender has dependents and equal to zero if not. Not a citizen is equal to one if the offender is not a US Citizen and equal to zero if not. Weapon is equal to one if the offense involves a weapon and equal to zero if not. Above median crim. hist. points is equal to one if the offender has a criminal history score above the median criminal history score for offenders and equal to zero if not. Above the median # of counts is equal to one if the offender has above the median number of total criminal counts for offenders and equal to zero if not. Column 9 tests for differences in bunching for cases with a "drug conspiracy" charge versus those without. Column 10 examines differences in bunching for offenders convicted in states with above/below the median fraction of black and Hispanic cases. The final column examines differences in bunching for offenders convicted in states with above/below the median level of racial animus. States above the median are: AL, AR, CT, DC, DE, FL, GA, IL, IN, KY, LA, MD, MI, MO, MS, NC, NJ, NV, NY, OH, OK, PA, RI, SC, TN, and WV. States below the median are: AK, AZ, CA, CO, HI, IA, ID, KS, MA, ME, MN, MT, ND, NE, NH, NM, OR, SD, TX, UT, VA, VT, WA, WI, and WY. Results are robust to removing cases in DC (see Table 2 notes). See Tables A14 and A18-19 for additional tests by district characteristics and Figure A12 for a plot of heterogeneity by presence of a conspiracy charge. The coefficients above are estimated from the following regression:

$$(11)\quad (280-290g)_{it} = \alpha_0 + \beta_1(\text{After2010}\times\text{W})_{it} + \beta_2(\text{After2010}\times\text{BH})_{it} + \beta_3(\text{After2010}\times\text{W}\times\text{Characteristic}^H)_{it} + \beta_4(\text{After2010}\times\text{BH}\times\text{Characteristic}^H)_{it} + \beta_5\text{Characteristic}^H_{it} + \beta_6\text{BH}_{it} + \beta_5(\text{Characteristic}^H\times\text{BH})_{it} + \epsilon_{it}$$

The results above are also robust to including all interactions from columns 1-9 in the same regression. In that regression, I reject $\beta_1 = \beta_2$ at the five percent level (p-value=0.0167).
*** p<0.01, ** p<0.05, * p<0.1

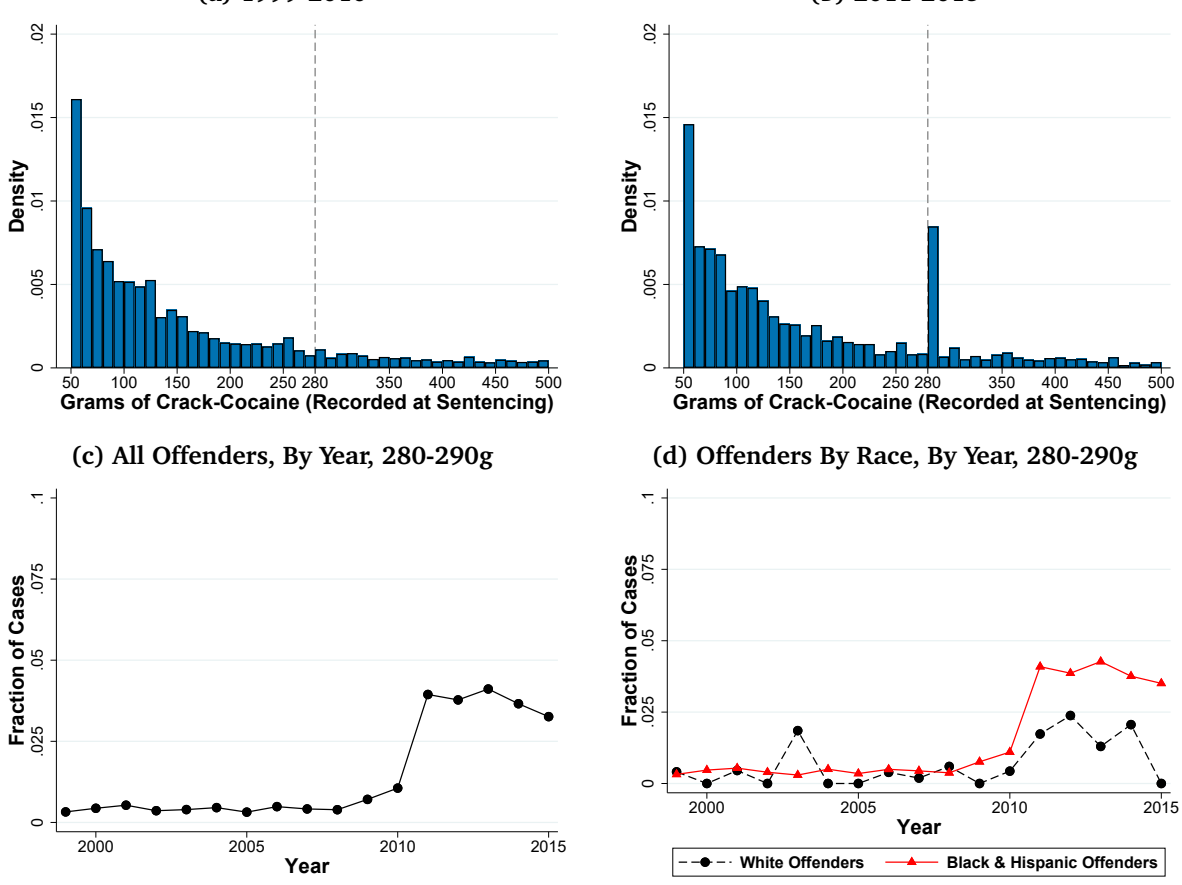
Table 9. Robustness Tests for Relationship between Racial Animus and the Racial Disparity Bunching at 280g

	Pr(280-290g)					
	(1)	(2)	(3)	(4)	(5)	(6)
After '10 x W x Above Med. Animus	0.0070	0.0059	0.0013	-0.0018		
	(0.0122)	(0.0134)	(0.0227)	(0.0225)		
After '10 x BH x Above Med. Animus	0.0191*	0.0220*	0.0320***	0.0318***		
	(0.0114)	(0.0125)	(0.0110)	(0.0108)		
After '10 x W x Continuous Animus					-0.0571	
					(0.0780)	
After '10 x BH x Continuous Animus					0.1073**	
					(0.0480)	
After '10 x IAT-Lawyers						-0.0084
						(0.0100)
After '10 x BH x IAT-Lawyers						0.0164
						(0.0108)
Constant	0.0058**	0.0008	-0.0230	-0.0225	-0.0047	0.0047
	(0.0022)	(0.0031)	(0.0286)	(0.0291)	(0.0322)	(0.0054)
Other Controls Included	None	Offender Controls	District Economic Controls	Offender + District Controls	Offender + District Controls	State x After 2010 x Race FEs
Observations	50,090	49,609	50,090	49,609	49,609	49,257

Notes. Standard errors clustered at the state level in parentheses for columns 1-5. Standard errors clustered at the district level are in parentheses for column 6. The estimates in this table are based on the USSC data. See Table 1 for notes on data construction. The first four columns examine differences in bunching for offenders convicted in states with above/below the median level of racial animus. Column 1 reports this result with no additional controls; column 2 introduces individual controls (college, male, age, criminal history, citizenship, dependents, state caseload, the average black-white gap in years sentenced at the district level pre-2010, and the average black-white gap in drug weights charged above 50g at the district level pre-2010) interacted with the after 2010 by race binary variables; column 3 introduces district controls for economic characteristics (median household income in 2016, non-white share of population in 2010, population density in 2010, fraction with college in 2010, poor share in 2010, log of wage growth for high school graduates, black-white and Hispanic-white differences in incarceration and income conditional on parent income rank at the 25th percentile, job density in 2013, and annual job growth from 2004-2013) interacted with the after 2010 by race binary variables; column 4 combines all controls from columns 2-3. Column 5 examines the relationship between animus and bunching using the continuous measure of animus from Google Trends (scaled from 0 to 1). Column 6 introduces a district level measure of animus, the implicit association test scores for lawyers (and other legal-service workers) aggregated to the district level. Since the measure is at the district level, I include state fixed effects interacted with the after 2010 by race binary variables. The estimate is identified from within-state variation in the IAT-animus measure, and the p-value on the estimate is 0.13. The IAT measure is scaled to the median difference between the minimum and maximum score in states, meaning a one unit increase is approximately equivalent to moving from the minimum score in a state to the maximum score. The relationship between state-level animus and bunching in column 1 is stronger after removing cases in DC, which has a fundamentally different prosecution system that relies more heavily on federal court, is very different demographically, and is below the median on the measure of state-level racial animus. This is partially addressed by controlling for offender characteristics in column 2 and fully addressed (i.e. estimates are the same with and without DC) by controlling for district characteristics in column 3. The results in column 1 are also robust to estimating separately for states that seceded during the US Civil War and states that did not.

*** p<0.01, ** p<0.05, * p<0.1

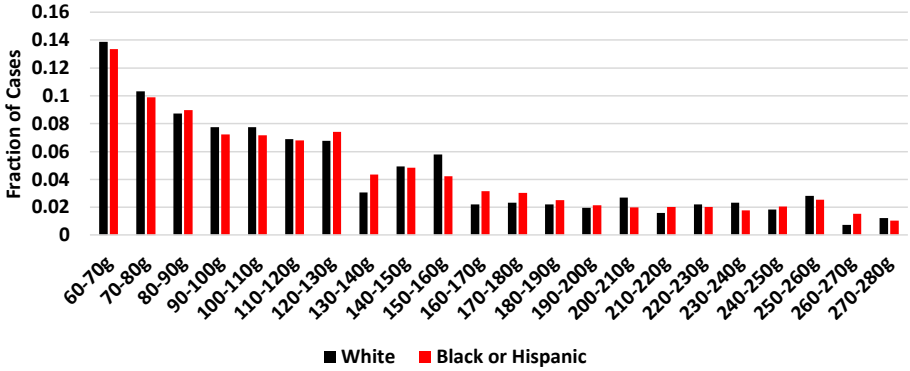
Figure 1. Changing Distribution of Drug Amounts Around 280g Pre- and Post-2010.



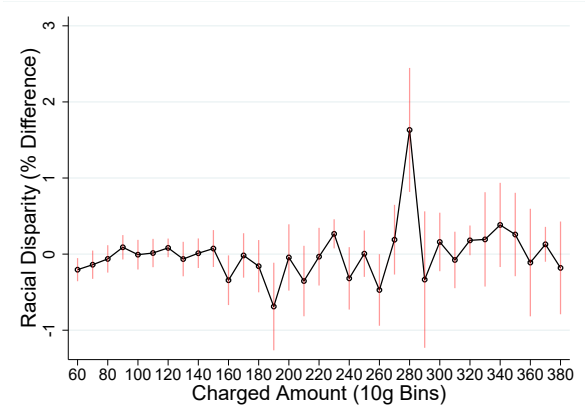
Notes. Panels (a) and (b) plot the distribution of drug amounts recorded in federal crack-cocaine sentences starting at (and including) 50 grams and ending at 500 grams for 1999-2010 (when the mandatory minimum threshold was 50g) and 2011-2015 (when it was 280g). The amount recorded at sentencing is not necessarily equal to the amount seized and the ways in which it can differ from the amount seized are discussed in detail in Section II.A (see Online Appendix B of Tuttle (2022) for further detail). Note the y-axis in these two figures is density not fraction or percent–this means the height of the bars is scaled such that the sum of the areas is equal to one. Panels (c) and (d) display the fraction of crack-cocaine cases with 280-290g by year, in general and by race. The denominator in panel (c) is all crack-cocaine cases under 1000g. The denominator in panel (d) is all crack-cocaine cases under 1000g, by race. Histograms showing the full density from 0-500g are in Figures A3a-b. Figures 1c-d with confidence intervals are in Figures A3c-d. See Figures A3e-f for alternative ways of visualizing bunching at 280g. These figures are based on the USSC data.

Figure 2. Testing for Conditional Racial Disparity in Bunching

(a) Distribution of Pre-2010 Charged Amount by Race, 60-280g

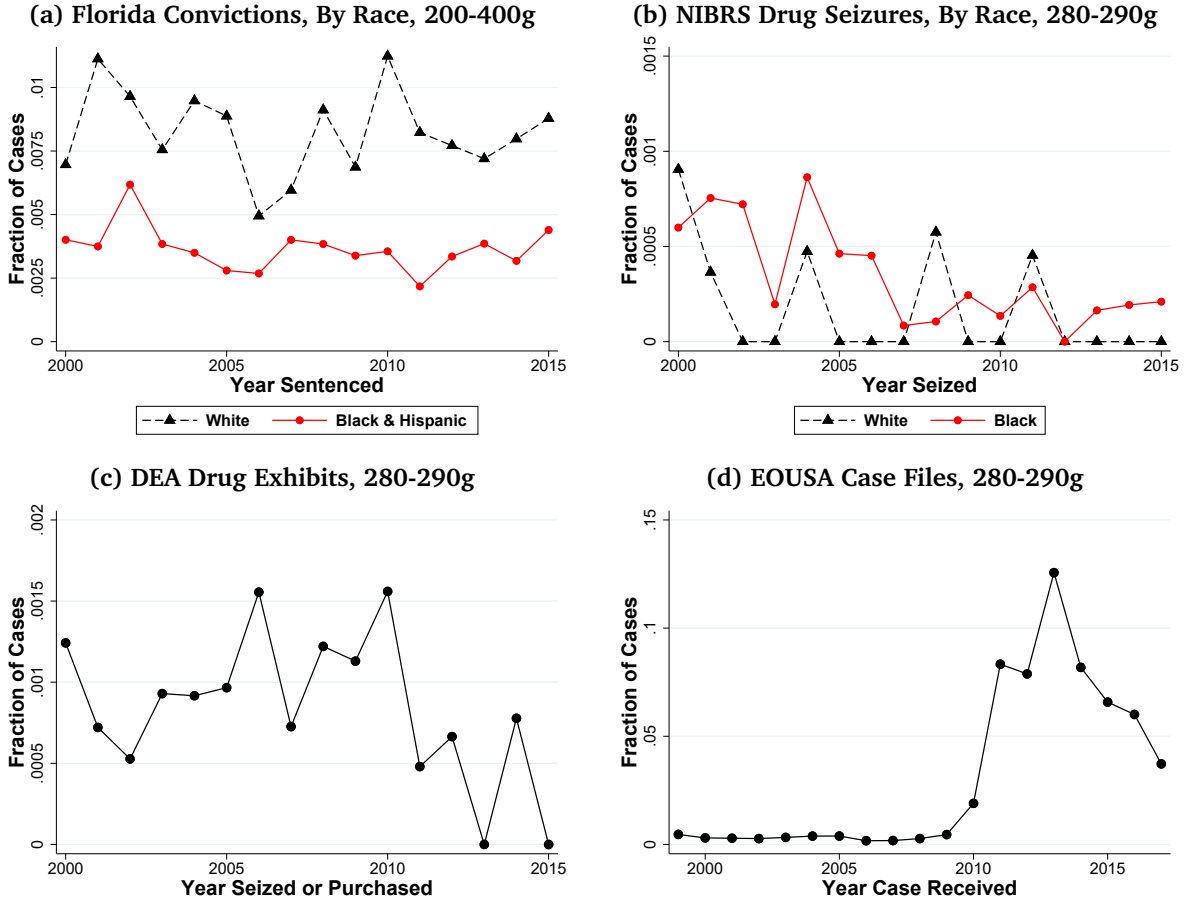


(b) Shifting from 60-380g by Race



Notes. Panel (a) plots the distribution of charged amounts pre-2010 from 60-280g. A Kolmogorov-Smirnov test of the equality of the distributions by race fails to reject the null that the distributions are equal (p-value=0.792). See Figures A5-6 for this distribution split by criminal history, predicted sentence, and state-level racial animus. Panel (b) plots the coefficient δ^X for each 10g bin starting at X divided by the share of cases in each 10g bin:(12) $(\text{Charged X-Yg})_{it} = \alpha + \delta^X(\text{After2010}\times\text{BlackOrHispanic})_{it} + \gamma\text{After2010}_{it} + \lambda\text{BlackOrHispanic}_i + X_{it} + \epsilon_{it}$ Eqn. (12) includes a set of controls such as age, college attainment, criminal history score, etc. to increase precision. Results are robust to removing cases in DC (see Table 2 notes). See Figure A7f for a version of this figure without controls and Figure A7e for a version of this figure in which controls are also interacted with the After2010 variable. The plot shows these estimates for amounts from 0-380g, at higher amounts the estimates are more noisy. Figure A7d shows the estimates up to 1000g. The coefficients at 160-170g, 190-200g, 210-220g, 240-250g, and 260-270g are statistically significant or marginally significant. The disparity in the overall decrease from 160-280g is also statistically significant. These figures are based on the USSC data.

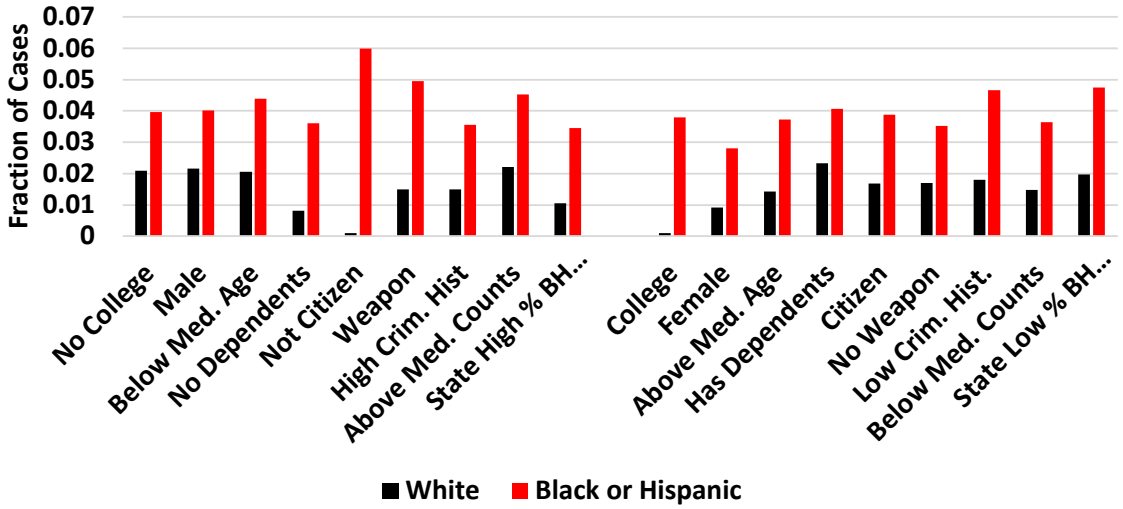
Figure 3. Changing Fraction of Cases at Various Stages of Criminal Justice System



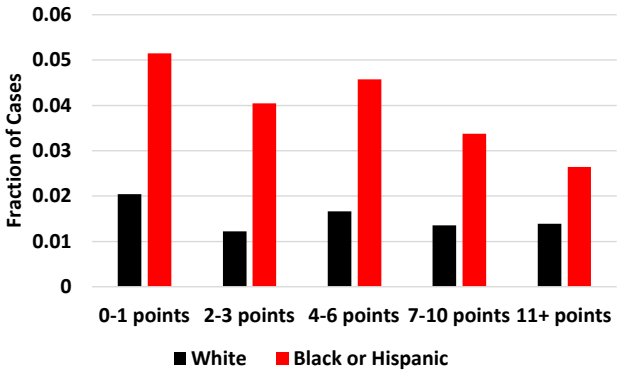
Notes. Please note the different y-axis scales, particularly in the case of panels (b) and (c). Panel (a) plots the fraction of cocaine offenses that have a range from 200-400g in FL state prison from 2000-2015, by race. The denominators are all cocaine offenses in FL, by race. Panel (b) plots the fraction of crack-cocaine drug seizures made by local police departments and recorded as 280-290g from 2000-2015, by race. Panel (c) plots the fraction of cocaine drug exhibits sent to DEA laboratories and recorded as 280-290g from 2000-2015 (the DEA data does not include race). The denominator is all cocaine exhibits in the DEA STRIDE data. Results are similar if limited to "cocaine hydrochloride" or "cocaine base." Panel (d) plots the fraction of crack-cocaine cases recorded as 280-290g in the EOUSA caseload data (the EOUSA data does not include race). The denominator is all crack-cocaine cases in the EOUSA data with non-missing drug quantities. The EOUSA data contains many more missing values than the USSC data. Imputing missing drug weights as zero does not fundamentally change the results. See Online Appendix E of Tuttle (2022) for further discussion and analysis of missing values in the EOUSA data. See Figure A8 for related figures.

Figure 4. Racial Disparity in Bunching at 280g Within Various Groups

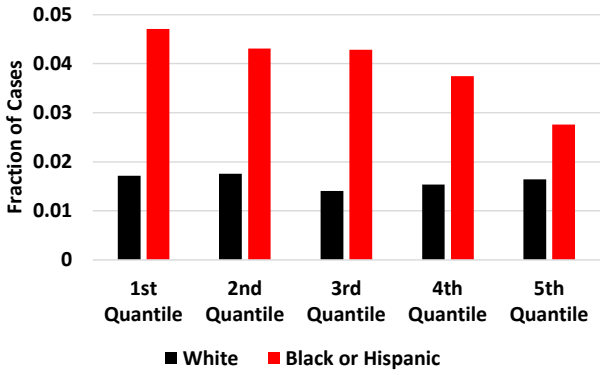
(a) Offender/Offense Characteristics



(b) Criminal History Score



(c) Predicted Sentence



Notes. Panel (a) plots the fraction of cases bunched at 280-290g after 2010 by race and within the following categories: college attainment, sex, above/below median age, whether the offender has dependents, citizenship, whether a weapon is involved in the offense, above/below median criminal history score, whether there is more than one count charged in the case, and whether the state where the offender is convicted is above/below the median in terms of fraction of black and Hispanic cases. Every within-category difference by race is significant at the 10 percent level and most are significant at the 5 percent level. Panel (b) plots the fraction of cases bunched at 280-290g after 2010 by race and within narrower groups of criminal history score that roughly correspond to quantiles. The within-category difference by race is significant at the 5 percent level for the categories 0-1, 2-3, and 4-6. For the 7-10 points category, the p-value is 0.15, and for the 11+ points category, the p-value is 0.383. See Figure A12a for a plot of bunching by race along the full distribution of criminal history score. Panel (c) plots the fraction of cases bunched at 280-290g after 2010 by race and within quantiles of predicted sentence. Using the pre-2010 data, I regress sentence length on criminal history score, sex, citizenship, age, education, number of dependents, and district where the case is sentenced. I take the coefficients from that regression and apply them to offenders in the post-2010 data to calculate a predicted sentence for each person. Finally, I split that predicted sentence into quantiles. See Figure A12b-c for a version of this figure in which sentence is predicted using exogenous and endogenous factors. The mean sentence within each quantile is: 4.7 years, 6.3 years, 7.7 years, 9.2 years, and 10.5 years. The correlation between actual sentence and predicted sentence post-2010 is 0.35. Finally, the within-category difference by race is significant at the 10 percent level for the first three quantiles. For the fourth quantile, the p-value is 0.166, and for the fifth quantile, the p-value is 0.505. These figures are based on the USSC data. Results are robust to removing cases in DC (see Table 2 notes)