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

Cody Tuttle

October 2025

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.

*Affiliation: UT Austin, Department of Economics. Email: cody.tuttle@utexas.edu. I thank Melissa Kearney for her 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, Felipe Goncalves, 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, both in terms of direct care costs (US DOJ 2011) and potential earnings losses (see, e.g., Mueller-Smith 2015). 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. 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.

In this paper, I examine racial sentencing disparities and test the second explanation: that agents in the criminal justice system treat black and Hispanic defendants differently than similar white defendants. 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 have power over key decisions related to case dismissal, diversion, and charging (Tonry 2012; Mueller-Smith and Schnepel 2021). The 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. Such quantity thresholds are common in the US legal system and abroad (Drápal and Šoltés 2023). In US federal cases, legal rules about the determination of drug involvement give police and prosecutors power to influence the amount used in sentencing. This paper studies whether they respond to this sentencing threshold and whether their responses are racially disparate. Specifically, I test for an excess mass (i.e., bunching) of cases at and above the mandatory

minimum threshold and for differences in the concentration of cases by race.

With the Fair Sentencing Act (FSA) in 2010, the 10-year mandatory minimum threshold for crack-cocaine was increased from 50g (i.e., 50 grams) to 280g. Crack-cocaine 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 pre-2010.[1] Using this time variation, I implement a difference-in-bunching design where I first assume the pre-2010 distribution of drug amounts is a good counterfactual for the post-2010 distribution in the absence of the threshold change (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 FSA, which I show, in line with prior work, caused a net decline in sentences after 2010 (USSC 2015a; Bjerk 2017a). Rather, these results imply that police or prosecutors curtailed the effect of the FSA by increasing the drug amount charged for some defendants.[2] 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. In short, while the FSA achieved its stated goal of narrowing racial gaps due to the crack-powder disparity in mandatory minimum thresholds, I find that the increase in bunching at 280g dampened its effectiveness. This strategic response further offers a unique window into discretion and disparities in the criminal legal system.

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. 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 is responsible for the bunching of cases at 280g, and I find that it is a result of prosecutor decisions. Third, I explore whether the disparity can be attributed to discrimination from prosecutors–I consider several alternative explanations and provide suggestive evidence of taste-based discrimination.

[1] All other mandatory minimum thresholds are set at somewhat natural bunching points (e.g., 50g).

[2] 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.

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.

Next, using data on drug use and selling, state-level drug convictions, and drug seizures, I rule out the possibilities that the bunching is a result of offender responses to the FSA or strategic enforcement 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 (EOUSA). I also find that approximately 30% of prosecutors are responsible for the rise in cases with 280g after 2010. These results 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 increased the evidentiary standard necessary for facts that raise a defendant's exposure to mandatory minimums and tasked juries, rather than judges, with evaluating said evidence in the event of a trial (Bala 2015). The EOUSA data show that from 2011-2013, 9.1% of cases were recorded in the 280-290g range, whereas only 6.8% of cases were recorded as 280-290g from 2014-2016. Using a difference-in-discontinuities design, I show that the practice of bunching ballooned in the run up to *Alleyne*, and was partially reined in by the Supreme Court decision.

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. Notably, the comparison of outcome rates by race (i.e., a "benchmark test"), only identifies taste-based bias under strong assumptions on the independence of race and unobserved factors and the role of race in decision-maker cost functions (Canay, Mogstad, and Mountjoy 2024). Accordingly, 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 among observably similar offenders. Next, I consider racial differences in defense or in other costs of developing a case. Last, I consider statistical versus taste-based discrimination from prosecutors.

I show that the racial disparity in bunching can be mostly explained by a measure of state-level racial animus (Stephens-Davidowitz 2014).

Related Literature Broadly, this paper adds to an extensive literature in economics on racial disparities and discrimination in the criminal justice system, and contributes to the empirical literature on prosecutorial discretion.[3] 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 courts. Yang (2016) highlights the importance of prosecutor discretion in federal cases by studying their decisions in the face of court resource constraints. Agan, Doleac, and Harvey (2023) show that more lenient prosecutors can reduce recidivism. In this paper, I explore prosecutorial discretion and its implications for racial disparities. Yang (2015) and Fischman and Schanzenbach (2012) document a racially disparate increase in sentence lengths concentrated at mandatory minimums after federal judges are given more discretion over other aspects of sentencing. 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. Sloan (2024) 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).

Researching disparities and bias from prosecutors, particularly those in federal court or on felony cases, has proven difficult. Rehavi and Starr (2014) made a major advance by introducing linked data from arrest to federal sentencing that permitted a detailed selection on observables

[3]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 et al. 2021; Cox and Cunningham 2020; Feigenberg and Miller 2020; Hoekstra and Sloan 2022; Luh 2023; Sloan 2024. 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 2024; Harrington and Shafer 2024; Jordan 2023; Ouss and Stevenson 2022; Sloan 2024; Agan, Doleac, and Harvey 2023. Literature outside economics is discussed briefly in Section II.

approach. They found that black offenders receive harsher sentences than white offenders arrested for the same crime, and that this is driven by prosecutor discretion to bring a charge with a mandatory minimum. A key concern they highlight is that even within offense codes, there may be unobserved differences in criminal conduct. In general, research on prosecutors faces a few critical challenges: (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 high stakes federal cases, prosecutors are not randomly assigned.

In this paper, I contribute to the literature in several ways. First, I address challenges (i) and (ii) 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 new evidence on prosecutor discretion in federal court, and I show that the burden of this discretion falls disproportionately on black and Hispanic offenders. Second, I address challenge (iii) by using the distribution of amounts charged pre-2010 as an empirical counterfactual to the post-2010 distribution. 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. By documenting that the amount distributions were similar by race prior to 2010, I provide new evidence that the disparate treatment in this setting is not due to differential drug involvement. Third, I quantify the fraction of prosecutors using this discretion, and I show that it may be mitigated by increasing evidentiary standards. Finally, I provide evidence on why this disparity arises by ruling out many alternative mechanisms and documenting a robust correlation between the racial gap in bunching and state-level racial animus.

Along with work on racial disparities in police ticketing 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. 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 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 given this vast discretion uses it in practice.

II. Institutional Background and Prosecutor Objectives

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. 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.

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. However, 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. Bjerk (2017b) briefly discusses bunching in the distribution of drug amounts, but posits that bunching arises from negotiation downward by prosecutors and defendants. 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 et al. 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.

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. 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.

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. 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.

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 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."

Prosecutor Objectives and Responses to the Fair Sentencing Act The fields of economics, criminology, and law all admit self-interested and/or biased prosecutors. 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 the innocent (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 prosecutor decisions (Rehavi and Starr 2014; Sloan 2024).

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).

This work 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.[4] In practice, longer sentences are handed down in cases just above the threshold (see Section V.E). 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.[5]

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. See Figure A2 for an illustration and elaborated discussion of these changes. Conceptually, 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. Also, some cases that were not bunched before 2010 will be worth bunching after 2010. In Section V.A.2, I explore empirical tests of this conceptual framework.

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 [...]. People will change their conduct to take advantage of lacunae." Since then, such rules have been implemented, but researchers have paid scant

[4]Note, these departures do reduce sentence length when applied, but to be eligible for these departures, defendants must cooperate with the government, which is itself costly.

[5]The cost of introducing new evidence, especially as that evidence grows weaker, can include costs of building the case, reputational costs, 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.

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 examine the role of prosecutors in bunching at 280g. Figure A1 shows a simplified timeline from arrest to sentencing and describes how the data I use are related to each step.

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.[6] 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 7.6% of offenders in this sample are labeled as white ($n = 3{,}798$), 9.2% as Hispanic ($n = 4{,}615$), and 83.3% as black ($n = 41{,}860$). Table 1 summarizes additional offender characteristics.

I restrict these data to cases in which the amount of drugs is non-missing and is not recorded as a range. Approximately 18% 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. However, 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. The notes of Table A1a discuss data construction and other minor restrictions in more detail.

B. Additional Data

I incorporate several other datasets to understand the source of the bunching in drug cases.

[6]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.

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 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. Sentence (in years) excludes life sentences and sentences less than 1 month, the by-period observation counts are 39,504 and 8,365, respectively, for this variable. Data construction is discussed in more detail in Table A1a.

Ancillary Records from State-level Cases, Drug Seizures, and Drug Use Surveys To study potential shifting from sub-federal to federal courts, I use state-level offense records from Florida's inmate database. To examine law enforcement responses to the FSA, I use drug seizure records from the National Incident Based Reporting System (NIBRS) Property Segment as well as the DEA System to Retrieve Information from Drug Evidence (STRIDE). I also use drug seizure records to analyze potential offender responses to the FSA. Finally, I further explore offender responses using survey-based measures of drug use and selling from the National

Survey on Drug Use and Health (NSDUH). Summary statistics and notes are in Table A1a.

Executive Office of the US Attorney (EOUSA), Caseload Data 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 software. 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. I use these data to explore the role of prosecutor discretion in bunching at 280g. Summary statistics are reported in Table A1b.

Google Trends Data on Racial Animus from Stephens-Davidowitz (2014) To measure racial animus at the state level, I use data introduced by Stephens-Davidowitz (2014). Stephens-Davidowitz uses Google Trends data from 2004-2007 and measures 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. 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 explore the role of prosecutorial discretion in causing 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 two goals. I reserve the discussion of prosecutor discretion for Section V.B and 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.

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 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. After2010_{it} is equal to one if the offender 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 280-290g range as a result of being sentenced after the threshold 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.[7] To understand where the excess mass at 280-290g might come from (i.e., where the post-2010 distribution has less mass relative to the pre-2010 distribution), I estimate models

[7]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 roughly similar for black and Hispanic offenders. 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).

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 this for 0-5g, 5-28g, 28-50g, 50-60g, 60-100g, 100-280g, 280-290g, 290-470g, 470-600g, 600-1000g, 0-280g, and 290-1000g.

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 *conditional* on drug involvement.

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. 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):

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

where β^{50} and δ^{50} represent shifts away from the 50g threshold for white offenders and black and Hispanic offenders, respectively. 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 offenders pre-2010, then equal distributions by race implies lower drug involvement for that group. Equal distributions by race only implies higher drug involvement for black and Hispanic offenders if the pre-2010 charged amounts were biased against white offenders.

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:

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

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 be shifted away from a narrow range than white offenders, that suggests the racial disparity in bunching at 280g is conditional on drug involvement.

V. Results

A. Main Results

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 over 2.5 times as likely to be charged with 280-290g after 2010 compared to white offenders. Figures 1a-c display graphical evidence of bunching at 280-290g and the racial disparity in that bunching.[8,9]

While there are nearly 4,000 white offenders in the full sample, Table 1 indicates there are only 544 white offenders after 2010. Given this sample size, one concern is that the disparity in

[8]While Figures 1a-b zoom in on the most relevant part of the quantity distribution, Figures A3a-b plot the histograms from 0-500g. Figures A3c-f present alternative ways to visualize this phenomenon.

[9]I focus on crack-cocaine because all other mandatory minimum thresholds are set at somewhat natural bunching points (e.g., 50g) that do not vary over time. However, Table A2 shows that disparities in bunching exist for all major drugs and discusses the difficulties in interpreting bunching in other drugs.

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

	Pr(280-290g Crack-Cocaine Recorded)	
	(1)	(2)
After 2010	0.0332***	
	(0.00204)	
After 2010 x White		0.0135**
		(0.0056)
After 2010 x Black or Hispanic		0.0344***
		(0.0021)
Constant	0.00487***	0.0031***
	(0.000343)	(0.0010)
P-value: W (White) = BH (Black or Hispanic)	-	0.0004
Observations	50,273	50,273

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." 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. Coefficients are estimated from eqn. (1) for Column 1 and eqn. (2) for Column 2. See Tables A3-A5 for additional robustness tests. *** p<0.01, ** p<0.05, * p<0.1

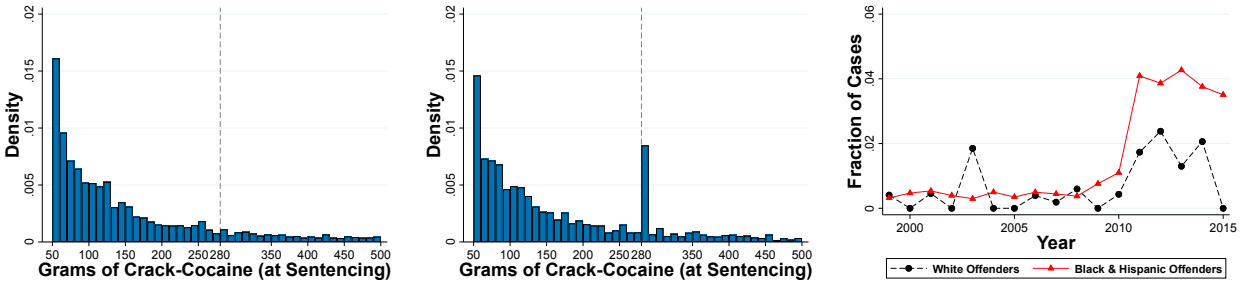
bunching has occurred by chance. To explore this possibility, I randomly assign white offenders to the 280-290g range based on the observed rates for black and Hispanic offenders both pre- and post-2010. Out of 500 placebo assignments, I only estimate a disparity as large (in absolute value) as the actual disparity in six cases. The median of these placebo estimates is 0.0004, the 90th percentile is 0.010, and the true disparity is 0.021. These results, shown in Figure A4a, demonstrate that the observed disparity from Table 2 is highly unlikely to occur under a null of no difference by race.

The main result is also robust to various sample restrictions; the inclusion of state fixed effects, time trends, state-specific time trends, and offender-level controls; 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 a range; and alternative methods of calculating standard errors. See Tables A3-A4 for these results. Notably, approximately 70% of the disparity remains even after controlling for state-by-post fixed effects and 60% remains with district-by-post fixed effects. In addition to these tests, Table A5 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.

Source of the Excess Mass at 280g To understand the reason for this bunching at 280g, I

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

(a) 1999-2010 (b) 2011-2015 (c) Share 280-290g



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 a few ways in which it can differ are discussed in Section II.A. Note the y-axis in these two figures is density, meaning the height of the bars is scaled such that the sum of the areas is equal to one. Panel (c) displays the fraction of crack-cocaine cases with 280-290g by year and by race. See Figures A3a-f for alternative ways of visualizing bunching at 280g. These figures are based on the USSC data.

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 above the threshold to secure longer or more certain sentences.[10]

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. The 95% confidence interval on this 80% estimate is [47.5%, 113.8%]. While the confidence interval on that share includes 100%, it bears emphasis that some offenders charged with 280-290g post-2010 may have been charged below 50g pre-2010 if there are fixed costs associated with bunching (e.g., a fixed cost of evidence-gathering).

I estimate the regressions in Table 3 by race in Table A6. The results are similar, particularly for black and Hispanic offenders, but less precise since they require cutting the already narrow ranges by race. Summing the coefficients in columns 4-6 of Table A6 for black and Hispanic offenders implies that the change in probability from 50g-280g can account for 88% of the increase in the 280-290g bin, with a 95% confidence interval ranging from [55.3%, 121.9%]. For white offenders, there is no net decrease in the 50-280g range; however, the 95% confidence interval on this increase is [-393%, 257%]. Despite the imprecision of this particular estimate,

[10]It is not possible 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. However, this analysis offers 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.

Table 3: "Missing Mass" in the Distribution of Drug Amounts Post-2010

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	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	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	-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	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 discussed in Section II.B is displayed in the row labeled "predicted change." Coefficients are estimated from eqn. (3) for each range listed in the column headers. Table A6 displays versions of missing mass results with time trend interactions and interactions by race. *** p<0.01, ** p<0.05, * p<0.1

Table A6 shows that the rise in bunching at 280-290g, the disparity in bunching, and the net decline from 50-280g for black and Hispanic offenders are statistically precise. Figures A5a-k plot the share of cases over time in each of these ranges.

To further examine changes in missing mass by race, I turn my focus to two broad ranges: 0-280g and 290-1000g. This analysis addresses the core question: can the increase in cases at 280-290g, and the disparity in that increase, be explained by shifts from above or below the new threshold? When examining these very broad ranges, I include linear time trends by race and district fixed effects to address long-run secular trends in these broad drug amounts. Table 4 shows that the rise in 280-290g cases for black and Hispanic offenders can be largely explained by a fall in cases below 280g. The results for white offenders are less precise; however, the decrease in the 0-280g range is larger in magnitude for black and Hispanic offenders than white offenders. This remains true when limiting to 2004-2015 to address differential trends in drug amount pre- and post-*Booker* (see columns 4-6 of Table 4). Table 4 also shows that, in general,

there is not a disproportionate 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 approximately as large as the disparity in whether a case is at 280-290g. Overall, these findings highlight that the cases charged with 280-290g post-2010 are plausibly drawn from the cases that would have been charged with 0-280g pre-2010, particularly for black and Hispanic offenders.

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, we should expect less bunching in trial cases where the plea bargaining channel is shut down. However, the degree of bunching is only heightened in trial cases ($\beta = 0.0674$ and $SE = 0.0129$), and the racial disparity is also larger. In fact, the only cases with 280-290g that go to trial are those of black and Hispanic offenders. It bears emphasis that the decision to go to trial is endogenously determined; still, these results establish the descriptive fact that bunching exists in trial cases and is not strictly an outcome agreed to in pleas. 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.

Table 4: "Missing Mass" in the Distribution of Drug Amounts Post-2010, by Race, with Race-Specific Trends and District FEs

	Pr(Below 280g)	Pr(280-290g)	Pr(Above 290g)	Pr(Below 280g)	Pr(280-290g)	Pr(Above 290g)
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010 x White	–0.0248	0.0146**	0.0102	–0.0021	0.0132*	–0.0111
	(0.0186)	(0.0064)	(0.0178)	(0.0238)	(0.0075)	(0.0229)
After 2010 x Black or Hispanic	–0.0574***	0.0318***	0.0256***	–0.0271***	0.0305***	–0.0034
	(0.0053)	(0.0024)	(0.0049)	(0.0067)	(0.0029)	(0.0062)
Constant	0.9424***	0.0051**	0.0525***	0.9751***	0.0022	0.0227
	(0.0090)	(0.0022)	(0.0087)	(0.0236)	(0.0045)	(0.0232)
P-value: W = BH	0.0927	0.0114	0.4047	0.3122	0.0312	0.7470
Sample	2000–15	2000–15	2000–15	2004–15	2004–15	2004–15
Observations	50,273	50,273	50,273	34,989	34,989	34,989

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. Coefficients are estimated from eqn. (3) for each range listed in the column headers. Columns (1)-(3) include all years in the sample period. Columns (4)-(6) focus on cases sentenced based on Guidelines year 2004 and beyond, to further account for secular trends in drug weight, particularly for white offenders, and for broad changes in charging pre- vs. post-*Booker*. Results are similar without district fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Estimating the Conditional Racial Disparity in Bunching at 280g The results above show that there is a racial disparity in bunching at 280g. 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. 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 2.1 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).

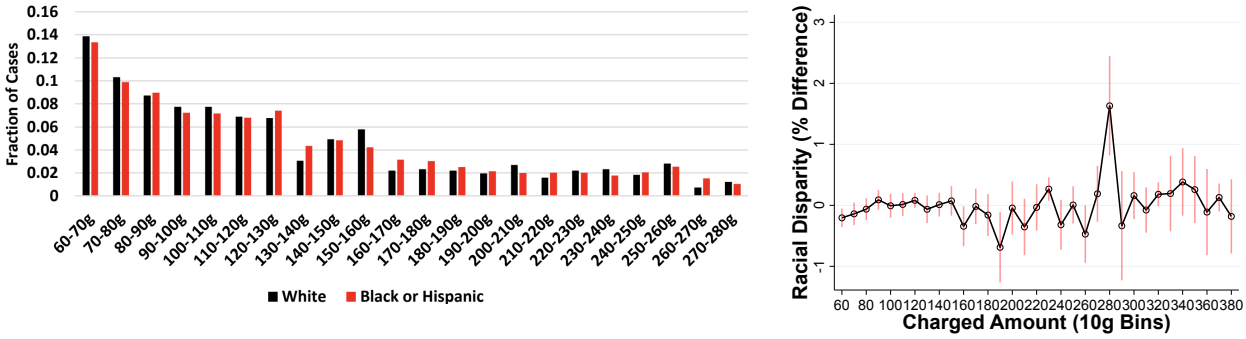
For the second part of the first test, I test whether the distributions of charged amounts from 60-280g are equal by race prior to 2010. Figure 2a plots the distributions by race, and they are very similar. A Kolmogorov-Smirnov test of equality fails to reject the null that the distributions are equal (p-value = 0.792).[11] 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.[12] 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 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 A8 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.

[11] Figure A6 breaks these distributions out by criminal history, predicted sentence, and state-level racial animus.

[12] Sentences do increase slightly above 50g, but the return to charging more from 60-280g is much lower pre-2010 vs. post-2010. For example, the sentence increase from 150-160g to 280-290g is over 6.5 times larger post-2010 than it is pre-2010. While there are thresholds above 50g that correspond to different Sentencing Guidelines, the Guidelines are advisory after 2005, and I do not observe bunching at Guidelines thresholds or a change in mass around those amounts once Guidelines thresholds change in 2010.

Figure 2: Testing for Conditional Racial Disparity in Bunching

(a) Pre-2010 Distributions by Race (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 Figure A6 for this distribution split by criminal history, predicted sentence, and state-level racial animus. Panel (b) plots the coefficient δ^X from equation (5) divided by the share of cases in each 10g bin. I include a set of controls such as age, college attainment, criminal history score, etc. to increase precision. 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.

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. There is a noisy decrease from 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. 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 shifts power to the jury and tightens evidentiary standards, further suggesting these cases may not be easily provable over 280g.

B. Prosecutorial Discretion

Ruling Out Earlier-Stage Responses Before turning to prosecutor responses, I first rule out other channels through which bunching at 280g may arise post-2010. The three main alternatives are that: (1) black and Hispanic offenders may respond differently than white offenders to the FSA, (2) state and local authorities could send more of their high drug weight cases to federal court after 2010, and (3) law enforcement agents may change investigatory tactics in response to the FSA. Tables A7 and A8 examine these three possibilities. First, I use drug seizure records and survey records to show that black and Hispanic offenders do not exhibit an increase in seized quantities, reported drug-use, or reported selling post-2010. Second, I use state-level records from Florida to show that its share of high drug weight cases does not change over this time period, despite bunching increasing in Florida's federal districts. Additionally, Figure A8e documents bunching among cases that solely involve federal law enforcement. Finally, I show that there is no bunching at 280g in drug seizure records, which suggests that discretion by law enforcement cannot explain the bunching in charged amounts after 2010.

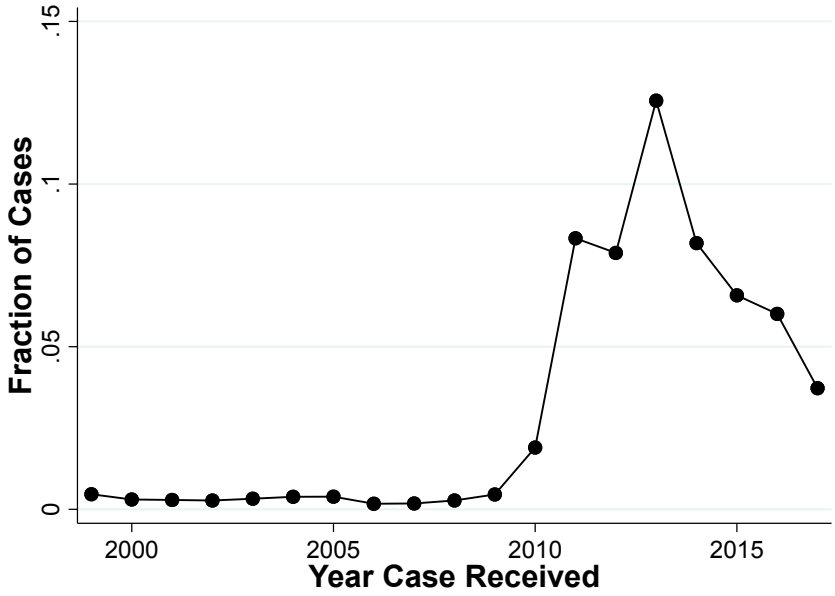
Bunching in Prosecutor Case Management Files The EOUSA provides case-level data extracted from their internal case management system. Using this data, I test for bunching in the quantity of drugs recorded in that system. Figure 3 shows a sharp increase in the fraction of cases recorded with 280-290g after 2010. 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.

In the EOUSA files, 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 yields an increase of about 2.4 percentage points after 2010, which is consistent with estimates from the sentencing data.[13]

The EOUSA data do not contain a field for race of the defendant or prosecutor. I impute defendant 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

[13]The main results below are robust to this missing value re-coding, see Table A9 and Figure A8g-h. For a more detailed discussion of missing values in the EOUSA data, see the notes of Table A1b.

Figure 3: Share of 280-290g Cases in EOUSA Case Files



Notes: This figure plots the fraction of crack-cocaine cases recorded as 280-290g in the EOUSA caseload data (note: 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 Table A1b notes for further discussion of missing values in the EOUSA data. See Figure A8 for related figures.

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.

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%. For this analysis, I limit the sample to attorneys with 10 or more drug cases after 2010.[14] Most attorneys will be

[14]All results are similar when using attorneys with 5+ or 15+ cases after 2010. See Table A10.

classified as "bunching" if they have at least one case in the 280-290g range after 2010.

The majority of these attorneys exhibit no bunching. Approximately 29.7% 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 0.025 and a 90% confidence interval of about 25.6-33.8%. Over 50% of bunching attorneys have two or more cases at 280-290g and over 30% 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.

The analysis above describes the share of attorneys who bunch at 280-290g and permits further comparisons between bunching and non-bunching attorneys. However, the small number of cases per attorney raises concerns that the uncertainty in these labels is too high to be informative about prosecutorial discretion, instead reflecting estimation error or the non-random assignment of cases. I address this concern using two complementary methods.

First, I implement a version of the outlier detection procedure from Ridgeway and MacDonald (2009), largely following Hoekstra and Sloan (2022). Specifically, I restrict the sample to prosecutors with 10+ cases post-2010 and regress the 280-290g indicator on a rich set of case characteristics: year fixed effects, case priority fixed effects, and binary controls for case complexity (e.g., multiple law enforcement agencies, gang involvement, multiple opposing counsel, etc.). I regress the residuals from this model on a set of prosecutor fixed effects, and use these estimated prosecutor effects to fit an empirical null distribution, to estimate the actual density, and calculate false discovery rates. Following Ridgeway and MacDonald (2009) and Hoekstra and Sloan (2022), I flag a prosecutor as an "outlier" if their false discovery rate is less than 0.5. Under this approach, approximately 22% of prosecutors are flagged as having a rate of bunching that is more extreme than that of the benchmark case (see Figure A9a). Using right-tail false discovery rates, approximately 27% of prosecutors are flagged as having a rate of bunching that is more extreme than that of the benchmark case.

Second, I apply a Bayesian shrinkage method using the discretized deconvolution approach of Goncalves and Mello (2021). After accounting for case-level observables, I estimate the true distribution of prosecutor effects via maximum likelihood, incorporating the standard errors of the estimated effects to account for noise. This approach suggests that approximately 30% ($SE = 0.042$) of prosecutors engage in bunching at 280-290g (see Figure A9b). Both methods

yield estimates of the share of bunching prosecutors that are consistent with the observed share.

This attorney-level bunching cannot be accounted for by district fixed effects. The within-district standard deviation in the 280-290g bunching metric is 0.147, the between-district standard deviation is similar at 0.149, and district fixed effects explain about 25% of the variance in the attorney-level bunching metric. Finally, 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. Table A9 explores this is in greater detail.

Further Evidence on Source of Excess Mass at 280g In Table 5, I estimate the likelihood a case is charged below 280g, with 280-290g, or above 290g for bunching versus non-bunching attorneys. This echoes the approach that Goncalves and Mello (2021) use to estimate bunching in speeding tickets. 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.

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. This assumption may be violated if cases are not randomly assigned to attorneys, an issue I probe further in the subsection on differences in case assignment. 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 increasing, rather than decreasing, 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

Table 5: 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	-0.1241*	0.1544***	-0.0303
	(0.0668)	(0.0438)	(0.0464)
Constant	0.8926***	0.0256***	0.0817*
	(0.0432)	(0.0074)	(0.0436)
Observations	974	974	974
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	-0.0803***	0.0597***	0.0207
	(0.0256)	(0.0175)	(0.0168)
Constant	0.9355***	0.0235**	0.0411***
	(0.0171)	(0.0106)	(0.0134)
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 and B are estimated from the following regression for each range: $(\text{Charged X-Yg})_i = \alpha_0 + \beta_1 \text{AttyBunchesAt}[X]g_i + \epsilon_i$. In panel A, $\text{AttyBunchesAt}[X]g$ 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 under that definition. 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. In panel B, $\text{AttyBunchesAt}[X]g$ 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 under that definition. These regressions are restricted to post-2010 cases and to prosecutors with 10+ cases pre-2010. See Tables A10-11 for additional robustness tests. *** p<0.01, ** p<0.05, * p<0.1

another actor in the 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. The figure notes 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. Given concerns about noise in the estimation of prosecutor-level bunching, these results that show prosecutor-level 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.

Differences in Case Assignment Since cases are not explicitly 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. In general, I do not find that bunching attorneys are involved in more complex cases (see Table A12 and Figure A12h). In addition, bunching at 280g is weakly correlated with prosecutor tenure in the EOUSA (see Figure A12g).[15] Recall that bunching prosecutors are also more likely to have cases at the 28g threshold after 2010 (see Table A12), and that bunching and non-bunching attorneys are equally likely to have cases above 290g (see Table 5). 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

[15] 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. Even in cases with only one defendant or with no witnesses, bunching attorneys are more likely to have cases at 280g than non-bunching attorneys.

power of the prosecution and defense, particularly in cases with evidence that would not satisfy the stricter evidentiary standard.[16,17]

Up to this point, I analyze bunching in the 280-290g range using a difference-in-bunching method that compares the pre-2010 and post-2010 drug distributions by race. This analysis rests on the assumption that the share of cases in the 280-290g range pre-2010 provides a valid counterfactual for the post-2010 share, absent the threshold change induced by the FSA. The share of cases in the 280-290g range is consistently near-zero from 2000-2010. This lends support to the underlying assumption and enables a characterization of bunching and the racial disparity in bunching over the full 2011-2015 period. However, such an assumption is difficult to justify pre- and post-*Alleyne*. *Alleyne* occurs in a period when the share of cases in the 280-290g range is already elevated and changing year-to-year. Put differently, it is unlikely that the share in the 280-290g range from 2011-2013 is a valid counterfactual for the share from 2013-2015, in the absence of *Alleyne*.

To address this, I adopt an alternative approach, which relies on the weaker assumption that case characteristics are continuous in a short window around the timing of *Alleyne*. Under this assumption, I estimate how prosecutors reacted to this decision by comparing the discontinuous 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 recorded in the EOUSA data for cases received immediately after *Alleyne*. Specifically, I estimate the effect of *Alleyne* as follows:

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

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**. Days_{it} is the number of days from June 17th that case **i** is received, and Year2013_{it}

[16]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.

[17]The decision may also change bargaining in cases with strong evidence, if judges and juries evaluate such evidence differently. The limited research on judge and jury decision-making finds that they agree in the vast majority of cases (Kalven and Zeisel 1967; Eisenberg et al. 2005). Notably, the share of cases above 290g does not appear to decrease after Alleyne ($\beta = 0.086$, $SE = 0.056$, based on +/- 150 bandwidth).

Table 6: Change in Bunching by Prosecutors after *Alleyne v. United States*

	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 regression discontinuity style model in eqn. (6). 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. See Table A13 for additional robustness tests. *** p<0.01, ** p<0.05, * p<0.1

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. δ_1 is the discontinuity that is specific to June 17, 2013–the date of the *Alleyne* decision.

Column 2 of Table 6 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.[18] 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*. Figure A11 illustrates the large discontinuity in the fraction of cases with 280-290g around June 17, 2013, displays the density of cases around this date, 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. This suggests that prosecutors were using discretion to build cases on evidence that was unlikely to pass "beyond a reasonable doubt" scrutiny from juries.

[18] 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.

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. Furthermore, 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 (see Table A16). 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 67% of those bunched at 280-290g fall outside of that range.[19]

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.

In terms of detecting discrimination, the approach in this paper falls into the broad class of "benchmark tests", in which one compares decision rates by race. As Canay, Mogstad, and Mountjoy (2024) discuss, benchmark tests are informative about taste-based bias under the assumptions that: (1) conditional on observables, race and unobserved factors that affect the decision are statistically independent and (2) the decision-maker does not consider race in the expected cost or outcome function. The setting in this paper partially bolsters assumption (1) by effectively examining a change in bunching among defendants with similar drug involvement, a crucial unobserved factor which affects prosecutor charging and which may be correlated with race. The following sections explore other non-race characteristics and the possibility of statistical discrimination. However, it bears emphasis that taste-based bias is only identified by the benchmark test when these two underlying assumptions are met.

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

[19] Didwania (2024) 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.

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 7, I show that the racial disparity in bunching exists even within many observably similar groups. Figure 4a examines the racial disparity in bunching within more narrow categories of criminal history score, showing that even among offenders with similar criminal history scores, there is a racial disparity in bunching. Finally, 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 4b, 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.

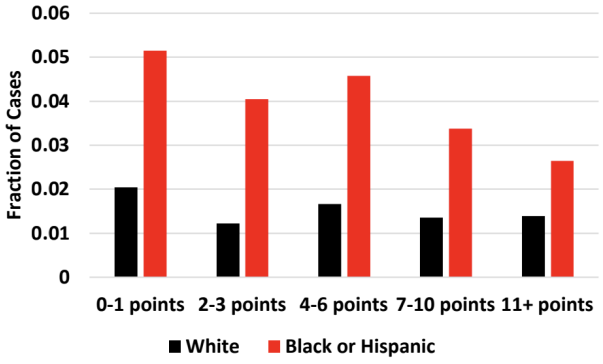
Table 7: Degree of Bunching Post-2010 by Race and Offender Characteristics

	Pr(280-290g)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
After '10 x 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 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 C	-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 C	-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 BH Cases	State Above Med. Racial Animus
P-value: W = BH	0.0195	0.0739	0.0415	0.0005	0.0010	0.0224	0.0048	0.0027	0.0000	0.1204	0.3799
P-value: W+C = BH+C	0.0000	0.0449	0.0038	0.0717	0.0003	0.0036	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

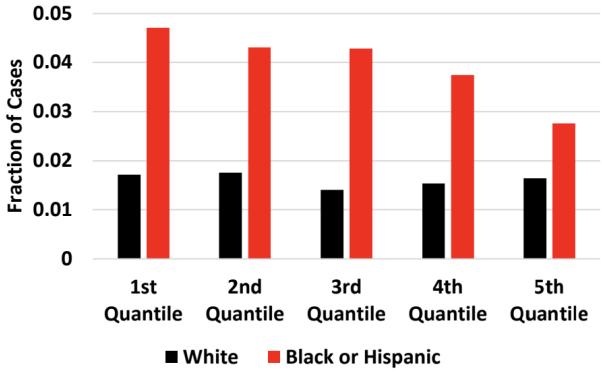
Notes: Robust standard errors in parentheses for columns 1-9. Standard errors clustered at the state level are in parentheses for columns 10 and 11. "Characteristic" or "Char." represents a binary variable that is an offender or case characteristic. The estimates in this table are based on the USSC data. 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)." The final column examines differences in bunching for offenders convicted in states with above/below the median level of racial animus. See Tables A14 and A19 for additional tests by district characteristics and Figure A12 for a plot of heterogeneity by presence of a conspiracy charge and state-level animus. The coefficients above are estimated from expanding eqn. (2) to include interactions between post-2010, offender race, and the listed characteristics. *** p<0.01, ** p<0.05, * p<0.1

Figure 4: Racial Disparity in Bunching at 280g Within Groups

(a) Criminal History Points



(b) Predicted Sentence



Notes: Panel (a) 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 (b) 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 0.34. 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.

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.[20] 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 7, column 9 and Figures A12d-e). Consistent with greater gang activity in the SIFCF, white offenders are, overall, more likely to face a conspiracy charge. This suggests that

[20]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 set of offenders does not change the main results. Likewise the presence of a weapons charge does not explain the racial disparity.

differences in gang participation by race do not explain the racial disparity in bunching.

Finally, I control for a proxy of individual drug involvement by linking the USSC cases to drug quantity info from NIBRS at the state-by-month-by-race level from NIBRS and from DEA STRIDE at the state-by-month level. This linkage is done based on month of sentencing in the USSC data, and sentencing in a federal case likely occurs long after the arrest. Thus, I control for lagged seizure quantities for up to 24 months prior to sentencing in these regressions. The first column of Table A17 reproduces the main result without controls. The second column includes 24 lagged drug seizure quantities from NIBRS (at the state, year-month, race level) and 24 lagged drug seizure quantities from DEA (at the state, year-month level). The increase in bunching at 280-290g and the racial disparity remains the same.

Although the factors considered above do not explain the disparity in bunching at 280g, there are, perhaps, other unobservable factors correlated with race that do. 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 alternative explanations. Ultimately, a compelling alternative explanation must not be captured by the detailed set of observables from the sentencing data, must address the evidence of similar drug involvement by race, and must comport with the decline in bunching observed after evidentiary standards are increased specifically in mandatory minimum cases.

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, since privately retained attorneys may be more likely to challenge cases narrowly bunched at 280g. The data do not include the offender's type of counsel in all years; however, from 1999-2002, black, Hispanic, and white crack-cocaine offenders are equally likely to be represented by private counsel (22.2% of white offenders, 21.9% of black offenders, and 21.9% of Hispanic offenders retain private counsel from '99-'02.)[21] 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. 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 use of bunching can be attributed to judicial preferences,

[21] Using data that includes all drug convictions (excluding marijuana) from the Federal Judicial Center, I find that private counsel retention rates are highly persistent within districts over time. The within-district correlation between mean private counsel retention in fiscal years 2000-2002 and in fiscal years 2000-2015 is 0.8.

since prosecutors may view bunching as less costly if they believe the judge will view the case favorably. 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, two other highly salient mandatory minimum thresholds for crack-cocaine (see Table A18). This suggests that bunching arises from prosecutors and is not merely a response to a judge-specific effect.[22] I also test whether district-level differences in costs of gathering evidence are related to the racial disparity in bunching. 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).[23]

Taste-based vs. Statistical Discrimination Lastly, I consider taste-based and statistical discrimination. These explanations are difficult to disentangle. Nevertheless, both are, ultimately, forms of racial discrimination, and 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)." Below, I discuss results which provide suggestive guidance about which form is at work in this setting.

Consider a model in which black and Hispanic defendants have higher true drug trafficking amounts on average, seized evidence is a noisy measure of true drug trafficking, the prosecutor's expectation over true drug trafficking amount determines bunching, and prosecutors within a district encounter the same composition of cases. In this model, statistical discrimination would imply that prosecutors within the same district should be equally likely to bunch cases at 280g. While I find that there is variation in the level of bunching across prosecutors within districts, this can be reconciled by relaxing the assumption that all prosecutors in a district see similar cases. More realistically, cases are not randomly assigned and some prosecutors may see more complex trafficking cases than others. Allowing nonrandom assignment of cases, the model above implies that accounting for other offender characteristics will decrease the racial disparity in bunching; however, I find that the racial gap exists within observably similar offender groups. Still, these results could be reconciled by a more detailed model of statistical discrimination.

To further investigate the possibility of statistical discrimination, I construct three measures

[22]Table A19 assuages concerns about juror bias because it shows that the disparity in bunching is not related to the share black in the district population (i.e., the potential jury pool).

[23]Another possibility is if black and Hispanic defendants are less likely to plead guilty, prosecutors may use mandatory minimums to induce a plea; however, the racial gap in bunching exists both in districts where black and Hispanic defendants are more likely to plead guilty pre-2010 and in districts where they are less likely.

that capture prosecutor experience and examine heterogeneity along these dimensions (see Figures A12g-i). First, I construct prosecutor tenure based on their first recorded crack-cocaine case in the EOUSA system. Second, for each case, I construct an index of case complexity for all previously handled crack-cocaine cases by the prosecutor in that case.[24] Third, for each case, I construct the mean weight involved in all previously handled cases by the prosecutor.

I find that the propensity to bunch does not change within the first five years of a prosecutor's tenure–prosecutors with 0-1 years of experience bunch at the same rate as prosecutors with 4-5 years of experience. However, the prevalence of bunching is lower for prosecutors with more than five years of experience, perhaps because attrition changes the sample composition. Using the index of prior case complexity, I find that bunching is unrelated to the complexity of the prosecutor's case history. Finally, prior weight charged by the prosecutor is unrelated to bunching in their current case across the first four quintiles of the prior weight distribution. However, prosecutors who have average prior weight in the top quintile are more likely to bunch at 280g after 2010. The similarity across most of the distribution with only a sharp difference at the top is consistent with the result that prosecutors who bunch one case are likely to bunch other cases. Altogether, these results provide further evidence against statistical discrimination: bunching is not strongly related to years of experience, the complexity of prior cases seen by the prosecutor, or, in general, the mean drug amount in the prosecutor's prior cases.

Another 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.

[24]The index is constructed by predicting drug weight in a case based on a regression of drug weight on district fixed effects; year fixed effects; variables indicating the case is a national priority, district priority, both, or neither; a variable indicating the case involves multiple agencies; a variable for gang involvement; variables indicating the case involves multiple opposing counsel or privately retained counsel; and a variable indicating the case has multiple court events recorded in the data. I take the average of this predicted drug weight for all prior cases as a measure of the complexity a prosecutor saw leading up to their current case.

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. Column 11 of Table 7 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 (see also Figure A12f).

Table A20 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. Column 5 estimates the relationship between bunching and the continuous measure of state-level animus from Google Trends. 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" (Xu et al. 2019). Since many states contain multiple federal districts, I include state fixed effects interacted with post-2010 by race, thus leveraging 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).

E. Sentencing Impacts

Sentencing Consequences of Bunching In order to understand the policy implications of bunching at 280g, I estimate the sentencing consequences of crossing the mandatory minimum threshold. Since mandatory minimums only constrain the minimum sentence, it is possible that being above the amount has no effect on actual sentencing. Judges could choose to treat defendants with 270g the same as defendants with 280g and sentence both to 10 years. I investigate this by employing an empirical strategy similar to Bjerk (2017b) and estimating:

$$Sentence_i = \alpha + \beta_1 Above280_i + \beta_2 Amount_i + \beta_3 (Above280 \times Amount)_i + \epsilon_i \quad (7)$$

where $Sentence_i$ is the sentence handed down for offender i, $Above280_i$ is equal to one if the offender is recorded with 280g or more of crack-cocaine and zero otherwise, and $Amount_i$ is equal to the offender's recorded drug quantity centered at 280g. I focus on cases sentenced after 2010. As long as the offenders who are bunched above the threshold are not negatively selected from the population just below the threshold, then β_1 will provide a conservative estimate of the sentencing penalty associated with crossing the mandatory minimum threshold after 2010. Since the bunching above 280g casts doubt on this assumption, I also estimate (1) for states

with low levels of bunching above 280g.

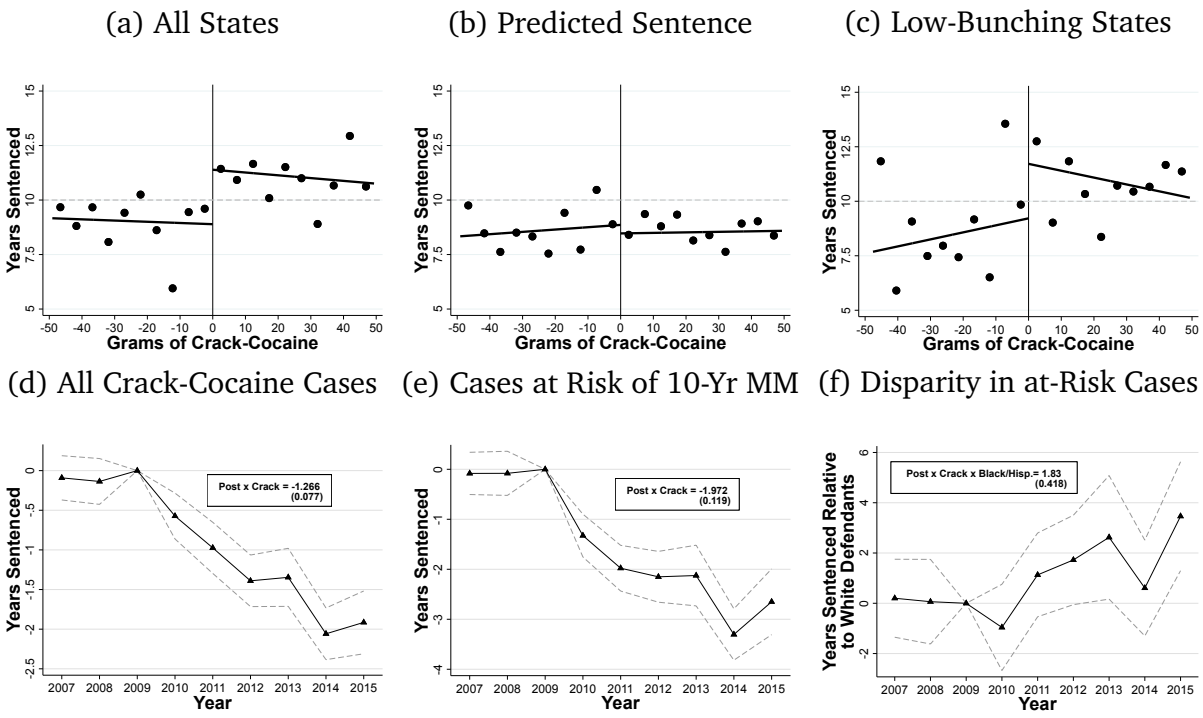
I find that bunching at 280g does have sentencing consequences. Offenders recorded with 270-280g after 2010 have a mean sentence of 9.6 years whereas offenders recorded with 280-290g after 2010 have a mean sentence of 11.4 years. Figure 5a plots sentencing outcomes by drug weight from 230-330g and the linear fit on each side of the 280g threshold for cases sentenced after 2010. The discontinuity (β_1) is the sentencing penalty from crossing the mandatory minimum threshold. Figure 5b shows that there is no discontinuity in predicted sentence, where sentence is predicted from a model using pre-2010 cases and several offender characteristics. Finally, Figure 5c plots actual sentence for the subset of cases sentenced in states that have low levels of bunching. Even in states where there is little manipulation around the threshold, there is a sentencing penalty of about 2 years.[25] This estimate assumes that an offender bunched at 280g would be charged with an amount just below 280g in the absence of the 280g threshold. However, Table 3 suggests that offenders bunched at 280g come from throughout the distribution below 280g. The average sentence after 2010 for offenders in the 50-280g range is 7.9 years.

Net Effects of the FSA The section above estimates the sentencing cost, to the defendant, of being charged just above the mandatory minimum threshold. Although bunching increases sentences for the affected defendants, the Fair Sentencing Act decreases mandatory minimum exposure for many defendants. How do these two forces interact to shape sentencing levels and racial gaps after 2010?

To answer this question about the overall effect of the Fair Sentencing Act, I build on Bjerk (2017a) and estimate a difference-in-differences specification that compares crack-cocaine cases to all other drug cases before and after the implementation of the Fair Sentencing Act in 2010. I find that, on average, sentences fall by about 1.3 years for crack-cocaine defendants. Furthermore, this decline is driven by defendants with amounts that would have exposed them to mandatory minimums pre-2010 but not post-2010. Among the affected group, sentences fall by about 2 years. Figures 5d-e show these results. Since black and Hispanic offenders are over-represented in crack-cocaine cases, this fall in sentencing narrows the overall racial gap in

[25] Bjerk (2017b), using data from fiscal years 2011-2012, employs a similar design and finds no sentencing penalty. I can fully replicate those results by limiting to those two fiscal years. The gist of the differences is that the USSC guidelines have changed over time, making sentencing for drug offenses below 280g more lenient and thus, giving the mandatory minimum more bite. Bjerk (2017b) makes the important point that the sentencing effect of the mandatory minimum is related to the overall federal sentencing structure that crack-cocaine offenders face, and this exercise confirms that point.

Figure 5: Sentencing Consequences of Bunching and Net Effect of the FSA



Notes: Figure 5a plots the average sentence (within each 5g bin) from 230-330g for cases sentenced after 2010 and a linear fit on each side of the 280g threshold. The estimated sentencing discontinuity is about 2.5 years (se=0.89). Figure 5b is the same plot but using predicted sentence from a model of sentencing and offender characteristics using pre-2010 data. There is no discontinuity in this figure, suggesting that, on net, offenders bunched at 280g are not negatively selected on characteristics that would increase sentence length in the absence of the threshold. Figure 5c is the same plot as 5a but limited to the subset of states that have low-levels of bunching. The estimated discontinuity is also about 2.5 years (se=1.85). These figures exclude life sentences and sentences less than 1 month, but results are robust to their inclusion. The coefficients described above are estimated from equation (7). Figure 5d plots coefficient estimates from a difference-in-differences event study comparing sentences for all crack-cocaine cases to sentences for other drug cases, before and after the Fair Sentencing Act (FSA). For this exercise, I focus on Guidelines years post-*Kimbrough*, a Supreme Court case that had broad sentencing implications for crack-cocaine cases. Assuming parallel trends, the pooled difference-in-differences estimate implies that the FSA decreased sentences of crack-cocaine cases by 1.3 years, on average. Figure 5e shows that this decrease is largely driven by a decrease in sentences for defendants charged with amounts greater than 50g, those most at risk of a 10-year mandatory minimum. The event study coefficients are estimated from: $\text{Sentence} = \alpha_0 + \sum_{k \neq 2009} \beta_k \mathbf{1}\{\text{Year}_t = k\} + \delta \text{Crack-Cocaine}_i + \sum_{k \neq 2009} \theta_k \mathbf{1}\{\text{Year}_t = k\} \times \text{Crack-Cocaine}_i + \varepsilon_{it}$ Figure 5f reconciles the results from Figures 5a and 5d. Although sentences decreased on net post-FSA, the strategic response of prosecutors dampened the potential decrease. For defendants at risk of a 10-year mandatory minimum, the racial gap in sentencing increased. Sentencing fell for both groups, but the disparate bunching response meant sentences did not fall as much for black and Hispanic offenders. The event study coefficients are estimated from a variant of the equation above that further interacts a "black or Hispanic" indicator with year fixed effects, crack-cocaine, and crack-by-year. These figures are based on the USSC data.

sentencing.

The combination of these analyses implies that although the FSA decreased sentences and overall disparities in sentencing, it would have decreased both *even more* were it not for the strategic response of prosecutors. To examine this directly in the difference-in-differences framework, I test how sentences changed post-2010 for black and Hispanic defendants with amounts above 50g relative to similarly positioned white defendants. In Figure 5f, I show that sentences increase relative to white defendants. In other words, while sentences fall for

both groups of crack-cocaine defendants, they fall more for crack-cocaine defendants who are white. The FSA succeeded in narrowing racial gaps due to the crack-powder disparity; however, strategic responses from prosecutors reduced its effectiveness. These strategic responses further offer new evidence on discretion and disparities in the criminal legal system.

Estimates of Sentencing Costs Finally, I use the estimates above to quantify: (1) the net benefit of the FSA and (2) the cost of prosecutor discretion. This exercise requires further assumptions about the cost of sentencing. I rely on estimates from Donohue (2009) and Mueller-Smith (2015) which suggest a per-year cost of approximately $60,000, accounting for operating costs and lost productivity. While estimates of post-release earnings losses vary, recent work continues to find earnings losses *during* incarceration and other financial consequences post-release (Garin et al. 2025; Humphries et al. 2025). Notably, operating costs remain high; in 2022, the Federal Bureau of Prisons estimated per capita operating costs of $138 per day, approximately $50,000 annually (BOP 2022).

Section E.2 shows that the FSA caused a 1.25 year decrease in sentencing, on average, in crack-cocaine cases. From 2011-2015, there were approximately 14,000 crack-cocaine defendants, implying a reduction of 17,500 sentence-years and a cost-reduction of just over one billion dollars. The vast majority of this benefit accrued to black defendants, as they made up approximately 85% of crack-cocaine defendants over this time period. On net, the FSA was successful at reducing sentencing and shrinking the racial gap caused by the 100:1 disparity in crack-powder mandatory minimum thresholds.

Despite its overall success, the FSA elicited strategic responses from prosecutors who pushed against the changes by increasing charged amounts in certain cases. This response allows me to study a clear case of prosecutor discretion and how it was applied differentially by race. 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. Although I focus on this specific empirical setting for identification benefits, I find that bunching disparities exist in other major drug types and attorneys who bunch at 280-290g also bunch at other mandatory minimum ranges pre- and post-2010. This suggests that bunching at 280-290g is indicative of more widespread discretion in the legal system, particularly for drug cases.

To give a conservative estimate, I assume that 3.4% and 1.3% of all federal drug cases from 1999-2015 were subject to similar discretion by race. In other words, I assume that prosecutors in the federal system exercised discretion to expose 3.4% of black and Hispanic drug defendants

and 1.3% of white drug defendants to an additional 2 years of sentencing. 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. I explore alternative assumptions in Table A21.

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 state-level cases 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 shifts power to the jury and requires stricter evidentiary standards for drug quantity evidence. This, in addition to numerous other tests discussed above, suggests that prosecutors are charging higher amounts to induce longer sentences.

Why do some prosecutors bunch black and Hispanic defendants at 280g more often than white defendants? The racial disparity cannot be explained by a rich set of observable individual characteristics or district characteristics. The failure of extensive case characteristics to explain the disparity, along with the fact that bunching is largely unrelated to prosecutor experience, casts doubt on some models of statistical discrimination. Alternatively, 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 have 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 evidence 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. I use these rates as a benchmark estimate for the extent of discretion in all federal drug cases from 1999-2015. Doing so 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.

Data availability statement

The data and code underlying this research are available on Zenodo at: https://doi.org/10.5281/zenodo.18343807

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* 138 (3): 1453-1505.

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. 2021. "The role of officer race and gender in police-civilian interactions in Chicago." *Science* 371 (6530): 696-702.

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 Review* 61 (4): 869.

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

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.

BOP. 2020. "Prison Safety." https://www.bop.gov/about/statistics/statistics_prison_safety.jsp.

BOP. 2022. "Federal Prison System Per Capita Costs." https://www.bop.gov/foia/docs/FY_22_Per_Capita.pdf.

Boylan, R. T. 2005. "What Do Prosecutors Maximize? Evidence from Careers of U.S. Attorneys." *American Law and Economics Review* 7 (2): 379-402.

Canay, I., M. Mogstad, and J. Mountjoy 2024. "On the Use of Outcome Tests for Detecting Bias in Decision Making." *The Review of Economic Studies* 91 (4): 2135–2167.

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."*Journal of Policy Analysis and Management* 40: 191-224.
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. 2024. "Charging Leniency and Federal Sentences." Forthcoming at *Journal of Legal Studies.*
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
Drápal, J. and M. Šoltés. 2023. "Sentencing decisions around quantity thresholds: theory and experiment." *Journal of Experimental Criminology* 20: 1323–1367.
Easterbrook, F. H. 1983. "Criminal Procedure as a Market System." *Journal of Legal Studies* 12 (2): 289–332.
Eisenberg, T., P. Hannaford-Agor, V. Hans, N. Waters, G. T. Munsterman, S. Schwab, and M. Wells. 2005. "Judge-Jury Agreement in Criminal Cases." *Journal of Empirical Legal Studies.* 2: 171-207.
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." *American Economic Journal: Economic Policy* 13 (2): 207-40.
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.
Garin, A., D. Koustas, C. McPherson, S. Norris, M. Pecenco, E. K. Rose, Y. Shem-Tov, and J. Weaver. 2025. "The Impact of Incarceration on Employment, Earnings, and Tax Filing." *Econometrica* 93 (2): 503-538.
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. 2024. "Brokers of Bias in the Criminal Justice System: Do Prosecutors Compound or Attenuate Racial Disparities Introduced by Police?" Working paper.
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: 389-414.
Humphries, J. E., C. Macaire, A. Ouss, M. T. Stevenson, W. van Dijk. 2025. "Revisiting the Lasting Impacts of Incarceration." Working paper.
Imbens, G., and K. Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933–959.
Jordan, A. 2023. "Plea Bargaining with Risk Aversion" Working paper.
Kalven, H. and H. Zeisel. 1967. "The American Jury." *Wash. & Lee L. Rev.* 24 (1): 158.
Kleven, H. J.. 2016. "Bunching." *Annual Review of Economics* 8: 435–464.
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.
Krumholz, S. 2019. "The Effect of District Attorneys on Local Criminal Justice Outcomes." Working paper.
Landes, W. M. 1971. "An Economic Analysis of the Courts." *Journal of Law and Economics* 14 (1): 61-107.
Luh, E. 2023. "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?" *American Economic Journal: Applied Economics*. 15 (3): 150-82
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. 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.
Ridgeway, G. and J. MacDonald 2009. "Doubly Robust Internal Benchmarking and False Discovery Rates for Detecting Racial Bias in Police Stops." *Journal of the American Statistical Association* 104 (486): 661-668.
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
Sklansky, D. A. 2018. "The Problems with Prosecutors." *Annual Review of Criminology* 1: 451–469.
Sloan, C. 2024. "Do Prosecutor and Defendant Race Pairings Matter? Evidence from Random Assignment." Working paper.
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.
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.