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

Cody Tuttle

October 19th, 2019

Job Market Paper

Please click here for the latest version

Abstract

I test for 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 discrimination. 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 almost entirely explained by a measure of state-level racial animus. These results shed light on the role of prosecutorial discretion and potentially racial discrimination as causes of racial disparities in sentencing.

JEL Classification: J15, K14, K41, K42

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

*University of Maryland, College Park, Department of Economics, 3114 Tydings Hall, College Park, MD 20740. Email: ctuttle@umd.edu. I thank Melissa Kearney for her continued guidance and support. For helpful comments, I thank David Arnold, Prateik Dalmia, Allan Drazen, Jessica Goldberg, John Haltiwanger, Judy Hellerstein, Brian Johnson, Ethan Kaplan, John Laub, Nolan Pope, Matthew Staiger, Lesley Turner, Mateo Uribe-Castro, Sergio Urzua, Stan Veuger, Riley Wilson, the participants of 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, and the Western Economics Association International Graduate Student Workshop. I thank Stephanie Rennane for sharing her bunching analysis code. Finally, I am grateful to Robert Breetz and several other unnamed attorneys for feedback and insights about the institutional setting.

I. Introduction

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 (United States Sentencing Commission (USSC) 2017). These added years are costly for society at large and for the people incarcerated. The Bureau of Prisons (BOP) estimates the direct care cost of incarcerating a person is about $11,000 (in 2015 dollars) per year (Department of Justice (DOJ) 2011). Mueller-Smith (2015) estimates an additional year in prison causes a 30 percent decrease in formal earnings post-release and significant lost wages while incarcerated. Even more, those incarcerated must confront serious physical and psychological costs of prison, in addition to the more intangible cost of their lost freedom (Haney 2001; The Hamilton Project 2016; BOP 2018). Due to racial sentencing disparities, these costs are disproportionately borne by black and Hispanic offenders.[1] For policy to confront these disparities, we must understand the root causes. One explanation for disparate sentences is that people of different races are different *upon entry* into the criminal justice system. Another explanation, however, is that *after entry* into the system, people are treated differently by race.

In this paper, I examine racial sentencing disparities and test the second explanation: that agents in the criminal justice system (police, prosecutors, judges, etc.) treat black and Hispanic defendants differently than similar white defendants.[2] To do this, I focus on federal crack-cocaine cases and the application of mandatory minimum sentences. Approximately 20 percent of all federal drug cases involve a crack-cocaine offense, and racial sentencing differences are particularly large in these cases. In 2016, black and Hispanic crack-cocaine offenders were sentenced to over 6 years, on average, compared to only 3.5 years for white crack-cocaine offenders (USSC 2017). In addition, the structure of mandatory minimum sentencing and recent changes in crack-cocaine mandatory minimums provide a unique opportunity to study discretion and racial disparities in the criminal justice system.

In federal drug trafficking cases, a mandatory minimum sentence is triggered if the drug trafficking crime involves an amount of drugs equal to or above a threshold amount. This sentencing cliff generates strong incentives for law enforcement agents. Legal rules about police sting operations and the type of evidence admissible in federal court give both police and prosecutors power to influence the amount used in sentencing. If police or prosecutors want to increase the likelihood of a harsh sentence, they can use their discretion to move the amount of drugs to the threshold amount or just above it. This paper studies whether police or prosecutors respond to this sentencing incentive and whether their responses are racially disparate. Specifically, I test for an excess mass (or bunching) of cases at and above the

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

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

mandatory minimum threshold (i.e. the use of discretion to increase the likelihood of a harsh sentence) and for differences in the excess mass by race (i.e. a racial disparity in the use of discretion).

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.[3] 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 since the new threshold is set at a point with zero bunching prior to 2010. All other mandatory minimum thresholds are set at somewhat natural bunching points (50g, 500g, 1000g) that do not vary over time.[4]

Using this time variation in the mandatory minimum threshold, I implement a difference-in-bunching design where I first assume the pre-2010 distribution of drug amounts is a good counterfactual for the post-2010 distribution (i.e. what the post-2010 distribution would look like with the pre-2010 thresholds) (Kleven 2016). I find the fraction of cases bunched at and above 280g increases after 2010, and that the increase is much larger for black and Hispanic offenders than for white offenders.[5] I then show further evidence that, under a few additional assumptions, this disparity in bunching at 280g is **conditional** on the observed drug trafficking of offenders and is not due to a difference in underlying observed drug trafficking by race.

To be clear, this is not intended as an evaluation of the FSA, which is likely responsible for a decline in sentences after 2010 (USSC 2015a). Rather, these results imply that police or prosecutors dampened the effect of the FSA by increasing the drug amount charged for some defendants. In addition, these results do not imply that the use of discretion or a racial disparity in the use of discretion began after 2010. Instead, I take the shift to 280g as an opportunity to detect these behaviors that are otherwise difficult to detect.

I use data at multiple stages in the criminal justice process to estimate who is responsible for the bunching at 280g. First, I use drug seizure records on quantities and prices and survey data on drug use and selling to show that offenders do not respond to the relaxed sentencing rules in a way that would induce this increase in cases at 280g (or the disproportionate increase by race). Second, since the bunching occurs in federal sentencing, it is possible that more cases with drug quantities at or above

[3]The FSA also shifted the 5-year threshold from 5g to 28g. I focus on the higher, 10-year mandatory minimum threshold for drugs in this paper. There are two reasons why I do not study bunching at 28g of crack-cocaine (the lower, 5-year mandatory minimum threshold) in detail. First, 28g is below the pre-2010 10-year mandatory minimum threshold of 50g–this yields incentives for prosecutors to shift cases that would have been charged both above 50g into the 28-50g range and cases that would have been charged below 50g into the 28-50g range. Second, estimating whether the racial disparity in bunching is conditional on underlying observed drug amount requires a range below the threshold that is not subject to strategic sentencing incentives. This is a reasonable assumption for the 60-280g range pre-2010, but would not be a reasonable assumption for the 6-28g range pre-2010 because those cases may be bunched at 50g.

[4]These amounts exhibit bunching in all drug types, even for drugs where they are not the relevant thresholds. I expect this bunching is due to a "round number" bias by police, prosecutors, offenders, etc.

[5]Note, I do not find evidence of bunching just below 280g for the drug amount used in sentencing. Moreover, comparing the pre-2010 and post-2010 distributions of crack-cocaine amounts suggests that these are cases that, had they been sentenced prior to 2010, would have been recorded below 280g.

280g are sent to federal court after 2010. I examine data on state-level drug convictions from Florida, and I do not find a shifting composition of cases after 2010. Third, local and federal law enforcement can influence the drug quantity involved in an offense by choosing amounts involved in sting operations. However, the data on drug seizures made by local and federal agencies do not show increased bunching at 280g after 2010.

Finally, prosecutors can legally influence the drug quantity involved in an offense because, according to the USSC Guidelines, 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 do 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, and that there is variation in prosecutor-level bunching both within and between districts. Prosecutors who bunch cases at 280g also have a high share of cases right above 28g after 2010 (the 5-year threshold post-2010) and a high share of cases above 50g prior to 2010 (the 10-year threshold pre-2010). Also, bunching above a mandatory minimum threshold persists across districts for prosecutors who switch districts. Moreover, when a "bunching" prosecutor switches into a new district, all other attorneys in that district increase their own bunching at mandatory minimums. 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 changed the evidentiary standard necessary for facts that raise a defendant's exposure to mandatory minimum sentencing (Bala 2015). Previously, prosecutors could present evidence on drug quantities to the presiding judge, and the judge would decide, based on the preponderance of evidence, whether the mandatory minimum applied. The Supreme Court ruling in *Alleyne* requires that prosecutors present this evidence to the jury, which evaluates it based on the stricter "beyond a reasonable doubt" standard. The case management data from the EOUSA show that from 2011-2013, approximately 9.1% of cases were recorded in the range of 280-290g. From 2014-2016, however, 6.8% of cases were recorded in the 280-290g range. Using a difference-in-discontinuities design, I show that the practice of bunching ballooned in the run up to *Alleyne*, and that this bunching was reined in by the Supreme Court decision (though it was not eliminated entirely). This suggests prosecutors were submitting evidence under the judicial fact-finding system that would not hold up under the scrutiny of a jury.

After documenting a racial disparity in bunching at 280g and studying the role of prosecutorial discretion in producing that disparity, I then explore whether the racial disparity can be attributed to discrimination. I introduce a simple model of prosecutor objectives and discuss four potential sources of the racial disparity. First, I explore the possibility that the racial differences in bunching at 280g are driven by another a factor correlated with race. I show that racial differences in bunching exist even among observably similar offenders. For example, the increase in cases at and above 280g for black and

Hispanic offenders with a college education is larger than the increase for white offenders with a college education. This is true for interactions with individual characteristics such as sex, age, criminal history, and other elements of the current offense. It is also true for interactions with district-level characteristics such as fraction of offenders who are white, pre-2010 plea rates, and pre-2010 fraction of cases declined. Race is a consistent factor in determining the amount of bunching at 280g after 2010.

Next, I test whether the disparity could be the result of racial differences in costs to the prosecutor of charging a defendant with 280g. Costs to the prosecutor are determined by defense attorneys, judges, potential juries, and other actors involved with the case. First, I show that there is no difference in type of defense attorney retained by race for federal crack-cocaine cases. Second, the increase in bunching at 280g is similar in districts with high versus low pre-2010 rates of private counsel retention. Third, I show that bunching at 280g is unrelated to judge race or political party and that, 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. Fourth, the increase in bunching at 280g is similar in districts with high versus low fractions of cases declined due to "weak evidence" or "lack of resources." These analyses suggest the racial disparity in bunching is not caused by racial differences in defense counsel, that bunching is not related to judges or judge characteristics, and that costs of developing a case are not a major determinant of the rise in bunching at 280g.

Finally, I consider statistical versus taste-based discrimination. I show that the racial disparity in bunching can be almost entirely explained by a measure of state-level racial animus based on Google search data developed by Stephens-Davidowitz (2014). In other words, black and Hispanic offenders convicted in states with higher levels of racial animus are more likely to be bunched at 280g than white offenders convicted in those states. In states with lower levels of racial animus, however, black, Hispanic, and white offenders are all equally likely to be bunched at 280g. The persistent racial differences even after controlling for and interacting race with observables, the within-district variation in prosecutor-level bunching, and the correlation between the racial disparity in bunching and state-level racial animus all support a model of discrimination in which the disproportionate use of discretion is a result of prosecutor tastes. Of course, a complicated model of statistical discrimination could incorporate those facts, and I cannot reject such a model.

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

Broadly, this paper adds to an extensive literature on racial disparities and discrimination in the criminal justice system (e.g. Knowles, Persico, and Todd 2001; Anwar and Fang 2006; Grogger and Ridgeway 2006; Antonovics and Knight 2009; Anwar, Bayer, and Hjalmarsson 2012; Rehavi and Starr 2014; Pfaff 2017; Arnold, Dobbie, and Yang 2018; West 2018; Sloan 2019). The vast majority of papers on this topic focus on racial bias from police officers and test for bias in two ways: (1) using a version of the outcome (or hit-rate) test proposed by Becker (1957) or (2) by documenting same-race versus other-race bias.

Along with recent work by Anbarci and Lee (2014) and Goncalves and Mello (2018), I implement a new test for racial bias in criminal justice that uses insights from the bunching literature.[6,7] Both Anbarci and Lee (2014) and Goncalves and Mello (2018) study the prevalence of police officers discounting speeding tickets by race. They show substantial bunching just below the point where the fine increases. Both papers argue that this is a result of officer leniency and that officers exhibit racial bias in their leniency.[8] I contribute to this literature by examining racial bias from prosecutors (a relatively understudied group), and by showing racial differences in bunching at the point where sentences increase.

This paper also contributes new evidence to the empirical literature on prosecutorial discretion and decision-making (e.g. Glaeser, Kessler, and Piehl 2000; Bjerk 2005; Boylan 2005; Shermer and Johnson 2010; Rehavi and Starr 2014; Yang 2017; Nyhan and Rehavi 2017; on defense attorneys: Agan, Freedman, and Owens 2018; Arora 2018; Carr and McClain 2018; Sloan 2019). Bjerk (2005), for example, finds that prosecutors are more likely to charge defendants with a misdemeanor if a felony charge would invoke a "three-strikes" sentence. Sloan (2019), using random assignment of prosecutors to cases in New York County, shows that being assigned to an opposite-race prosecutor increases a defendant's likelihood of conviction, particularly in property crime cases.

The most closely related work, Rehavi and Starr (2014), finds that black offenders receive harsher sentences than white offenders arrested for the same crime. Using linked data from US Marshals, US courts, and US federal sentencing, they show that this disparity is driven by prosecutorial discretion over initial charging decisions, in particular, the decision to bring a charge with a mandatory minimum sentence.[9] In this paper, I provide novel evidence that prosecutors are selectively harsh by race using a new source of identification–the sharp change in the crack-cocaine mandatory minimum threshold. I

[6]Note, my paper is not the first to acknowledge the existence of bunching in the amount of drugs recorded in US federal sentencing or the possibility that it could be used as a test of prosecutorial discretion and discrimination. However, this paper is the first, to my knowledge, to take advantage of the time variation in the crack-cocaine 10-year mandatory minimum threshold to isolate bunching that is solely due to the prosecutor. In addition, I examine data at multiple stages in the criminal justice process and conduct several additional empirical tests that all suggest bunching is due to prosecutorial discretion and negatively affects minority defendants. Related work in this area is discussed in more detail in Section II.A.

[7]Recently, economists have also studied bunching around cliffs and notches in test scores as evidence of manipulation in educational settings. See Diamond and Persson (2017) and Dee et al. (2017).

[8]Anbarci and Lee (2014) show that white officers discount more for white drivers and black officers discount more for black drivers. Goncalves and Mello (2018) demonstrate that only some officers practice this leniency and that those officers are, on average, more lenient toward white drivers than minority drivers.

[9]Rehavi and Starr (2014) do not focus on racial disparities in drug offenses due to data limitations.

argue that the sudden increase in cases just meeting that threshold is indicative of discretion, and that the burden of this discretion falls disproportionately on black and Hispanic offenders. Through a series of tests, I find that prosecutors are responsible for the increase of cases at 280g. In addition, I quantify the fraction of prosecutors exercising this type of discretion, and I show that this can be mitigated by increasing evidentiary standards.

Finally, the racial disparity in bunching at 280g has meaningful implications for the racial sentencing gap. Depending on the counterfactual sentence imputed for the affected offenders, bunching at 280g can account for 2-7 percent of the racial disparity in crack-cocaine sentences. A conservative estimate suggests that being bunched at 280g adds 1-2 years to an offender's sentence. Multiple estimates suggest the cost of incarceration (combining direct care costs and the cost of lost current and future wages for the offender) is approximately \$60,000 per person per year (Donohue 2009; Mueller-Smith 2015).[10] I find 3.6% of black and Hispanic crack-cocaine offenders are bunched at 280g after 2010 versus 1.2% of white crack-cocaine offenders. Assuming 3.6% and 1.2% of all drug cases from 1999-2015 were subject to similar discretion by race implies total costs of 1.3 billion dollars for black and Hispanic offenders versus 148 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 2,500 years sentenced for white offenders.

All of the calculations above are based on the amount of discretion and the disparity detected right at and above the 10-year mandatory minimum threshold for crack-cocaine. To the extent that prosecutors exercise similar discretion to push defendants just above 5-year mandatory minimum thresholds or exercise discretion in less obvious ways (pushing defendants far beyond thresholds, for example), the cost estimates will only be higher and the effect on racial sentencing differences will only be greater.

II. Institutional Background and Prosecutor Objectives

A. Institutional Background

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

Debate about federal mandatory minimum policy has overwhelmingly focused on the disparity between the threshold amounts for crack-cocaine and powder-cocaine. Prior to 2010, the threshold for the crack-cocaine 10-year mandatory minimum was 50 grams whereas the 10-year threshold amount for powder-cocaine was 5000g, a 100-to-1 disparity. In part due to the recommendations of the 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. The upper threshold was changed from 50g to 280g, and offenders

[10] The majority of inmates in the Survey of Inmates in Federal Corrections (2004) report earning formal wages in the month before arrest.

sentenced after the Fair Sentencing Act are subject to the new threshold.[11] 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.[12] Bjerk (2017) briefly discusses bunching in the distribution of drug amounts, but posits that bunching arises from negotiation downward by prosecutors and defendants.[13] A 2015 Bureau of Justice Statistics (BJS) working paper on federal sentencing disparities also investigates the idea that prosecutors could "game" the drug weight sentencing guidelines (Rhodes, Kling, Luallen, and Dyous 2015). That paper 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.[14]

I depart from previous work in several ways. First, I show that excess mass at the threshold comes from cases below the threshold rather than above it. I also show that the bunching is more pronounced in trial cases, which suggests that drug amounts are being moved above the cutoff and not negotiated down to it. Second, I take advantage of the time variation in the crack-cocaine 10-year mandatory minimum threshold to isolate bunching that is solely due to prosecutor choices. Finally, I examine data at multiple stages in the criminal justice process and conduct several empirical tests that all suggest prosecutorial discretion negatively affects minority defendants.

2. Procedural Background

In Figure A1, I illustrate a simplified timeline from arrest to sentencing. Arrests are made by local or federal police, and after arrest, cases are handled by state or federal prosecutors. Prosecutors decide whether to try the case in court. 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. A case tried 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 defendant's sentence. In

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

[12] In concurrent work, Knorre (2017) finds evidence of bunching in reported drug amounts from Russian police. Knorre does not investigate potential discriminatory behavior or the consequences of the observed bunching.

[13] Since Bjerk's paper focuses on sentencing consequences of mandatory minimums for all drug types, he does not empirically investigate the cause of the observed bunching in crack-cocaine offenses. In addition, he does not compare outcomes before and after the Fair Sentencing Act of 2010.

[14] The working paper is an extensive and excellent treatment of sentencing disparities. In that light, it is reasonable that the authors did not do a "deep dive" on this "bunching" test, which is a small piece of the broader paper.

2015, approximately 70% of drug arrests referred to federal prosecutors were prosecuted and 90% of those prosecuted ended in a conviction (BJS 2016). The drug quantity used in sentencing can be influenced at many of these stages. Below, I describe the legal discretion that police and prosecutors have over the drug quantity.

First, police can influence drug amounts by choosing the amount of drugs involved in "reverse sting" operations (operations in which agents will sell drugs to offenders) or by extending traditional sting operations (operations in which agents will buy drugs from offenders) until the total transacted amount is above the threshold (Honold 2014). Outside of these two levers, it is unlikely that law enforcement agents across multiple agencies could systematically manipulate drug amounts since evidentiary protocols require the precise logging and controlled storage of evidence.

Second, prosecutors can influence drug amounts because mandatory minimum sentencing is determined by the 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 "historical weight," the amount of drugs a defendant is responsible for outside of the actual drugs seized (Lynch 2016). In addition, mandatory minimums also apply to drug trafficking conspiracy crimes 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) specifically state, "Types and quantities of drugs not specified in the count of conviction may be considered in determining the offense level. 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 controlled substance."

Criminologist Mona Lynch has compiled compelling qualitative evidence about the reach of federal sentencing guidelines in her book *Hard Bargains*. Lynch finds that prosecutors use informants to establish "relevant" quantities, and she interviews a prosecutor about how relevant quantities can be established: "The actual heroin sales directly tied to Mr. Samuels and his son were of 1g and 4g, respectively; the rest was arrived at on the mere say-so of confidential informants. [...] She told me that she could have established enough historical weight, through those (conspirators) she had 'flipped,' to get Mr. Samuels to at least a ten-year mandatory minimum sentence, if not more."

In Section V.C, I examine data from a national survey on drug use/selling, state-level convictions, local police agencies, the Drug Enforcement Administration, and the Executive Office of the US Attorney to estimate the source of the bunching at 280g. I also conduct several tests in Sections V.C-V.E to rule out alternative explanations related to the role of offenders, state courts, police, defense attorneys, probation officers, and judges. Ultimately, I find evidence that prosecutorial discretion leads to bunching at 280g in the case of drug trafficking.

B. Prosecutor Objectives

Prosecutors have discretion over the drug quantity charged in federal drug trafficking cases. In addition, the data suggests prosecutors exercise this discretion and that they exercise it differentially by race. In this section, I discuss the literature on prosecutor objectives from the fields of economics, criminology, and law–all of which admit self-interested and/or biased prosecutors.

Then, in light of the literature on prosecutor objectives, I discuss how sentence-maximizing prosecutors may respond to the Fair Sentencing Act. 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.[15] Although this conceptual discussion describes prosecutor objectives as homogenous, I ultimately find that only a subset of prosecutors behave in this way.

1. Related Literature

Since the 1970s, economists have produced several theoretical models of plea-bargaining based on prosecutor objective functions. This work began with the canonical economic model of the courts from Landes (1971), which assumes that prosecutors maximize the expected sum of sentences subject to resource constraints. Following Landes (1971), several papers emerged modeling resource-constrained prosecutors trying to achieve an ideal punishment for guilty parties and no punishment for innocent parties (Grossman and Katz 1983; Reiganum 1988; Bjerk 2007; and Baker and Mezzetti 2011).

Empirical work finds that prosecutors are, in part, career-focused (Glaeser, Kessler, and Piehl 2000; Boylan 2005). Boylan (2005) shows that for US attorneys longer sentences are associated with positive career outcomes (appointed to a federal judgeship or hired by a large private firm). In addition, recent work demonstrates partisan bias (Nyhan and Rehavi 2017) and racial bias (Rehavi and Starr 2014; Sloan 2019) in prosecutorial decisions, suggesting that prosecutors may seek harsh punishments for some offenders and lenient punishments for others.

These findings that prosecutors can be self-interested and biased are echoed and often-times preceded by insights from criminologists and legal scholars.[16] Discussions of prosecutorial discretion in law reviews frequently note that career-oriented prosecutors focus on securing lengthy sentences or high conviction rates (Bibas 2004; Simon 2007; Barkow 2009; Sklansky 2017). Stuntz (2004) argues that prosecutors lean on harsh sentences to secure guilty pleas. He even specifically notes the usefulness of sentencing guidelines (e.g. mandatory minimums) in this regard: "plea bargains outside the law's

[15] Mandatory minimums also provide certainty about sentence length. Thus, in this context, prosecutors who desire certain sentences will behave similarly to prosecutors who desire long sentences.

[16] Officially, the EOUSA cites Berger v. United States, 295 U.S. 78 (1935) to describe the role of the US attorney as an agent "[...] whose interest, therefore, in a criminal prosecution is not that it shall win a case, but that justice shall be done. [...] the twofold aim of which that guilt shall not escape or innocence suffer. [...] It is as much his duty to refrain from improper methods [...] as it is to use every legitimate means to bring a just one." However, the quote offers a description of the prosecutorial ideal rather than the reality. In fact, the case in Berger v. United States, is itself a case about prosecutorial misconduct.

shadow depend on prosecutors' ability to make credible threats of severe post-trial sentences. Sentencing guidelines make it easy to issue those threats."

Finally, criminologists and political scientists have also documented prosecutorial bias along race, gender, and partisan lines (Spohn, Gruhl, and Welch 1987; Mustard 2001; Farrell 2003; Ulmer, Kurlychek, and Kramer 2007; Gordon 2009; Shermer and Johnson 2010; Fischman and Schanzenbach 2012; Ulmer, Painter-Davis, and Tinik 2014; Franklin and Henry 2019; King 2019). Fischman and Schanzenbach (2012) show that sentence lengths are concentrated at mandatory minimums, that this concentration grows when judges are given more discretion over other aspects of sentencing, and that the increase in bunching at mandatory minimum sentence lengths is especially large for black and Hispanic offenders. Farrell (2003) and Ulmer, Kurlychek, and Kramer (2007) both use state court data to show that black offenders are more likely to receive a mandatory minimum penalty than white offenders, even after conditioning on several aspects of the offense. Ulmer et al. (2007) conclude, "prosecutors have great influence through charging, sentence bargaining, and, in the case examined here, the application of mandatory minimums. [...] Too often, studies of sentencing and sentencing discretion focus on judges and leave out prosecutors, crucial players in the courtroom work groups."

2. Prosecutor Responses to the Fair Sentencing Act

This work from economics, criminology, and law suggests that prosecutors will value crossing the mandatory minimum threshold in drug cases (for sentence length and/or sentence certainty) and that they will value it differentially by race (due to racial bias). By law, cases above the mandatory minimum threshold must receive a sentence of at least five or ten years (increased certainty), and in practice, longer sentences are handed down in cases just above the threshold (increased sentence length; see Section V.B). Assuming that gathering new evidence to raise the drug quantity charged beyond the amount seized is costly and that the cost is increasing in the amount of new evidence gathered, these objectives yield predictions for how prosecutors will behave in the face of mandatory minimum thresholds and how they will behave when those thresholds change.

Prior to 2010, the mandatory minimum thresholds in federal court for crack-cocaine were 5g (for a five-year mandatory minimum sentence) and 50g (for a ten-year mandatory minimum sentence). After 2010, these thresholds shift to 28g and 280g. The shift in mandatory minimum thresholds after 2010 should lead to the following relative changes: (1) an increase in the density from 0-5g, (2) an ambiguous change in the density from 5-28g, (3) an increase in the density from 28-50g, (4) a decrease in the density from 50-280g, (5) an increase in the density from 280-290g, and (6) no change in the density above 290g. Note, for these ranges, and whenever ranges are listed, the upper bound of the range is not inclusive. See Figure A2 for an illustration of these changes.

These changes should occur because some cases worth bunching at 5g or 50g before 2010 will also be worth bunching at 28g or 280g after 2010 and some will no longer be worth it. Also, some cases that were not bunched at 5g or 50g before 2010 will be worth bunching at 28g or 280g after 2010. In Section V.E.1, I introduce a simple model of prosecutor objectives to motivate a discussion about the racial disparity in bunching. I use that model in Appendix C to formally discuss why the changes described above should occur. In Section V.A.3, I show that the empirical evidence is consistent with this simple conceptual model of prosecutor responses to the shifting thresholds.

This conceptual discussion and the empirical analysis that follows is rooted in broad ideas about prosecutor bias and prosecutors' desire for long sentences and/or certain sentences, but it also captures a specific phenomenon that has received some attention in law and criminology–federal prosecutors using sentencing guidelines and mandatory minimums to secure guilty pleas or harsh sentences (Stuntz 2004; Honold 2014; Lynch 2016). As noted in the previous section, Honold (2014) and Lynch (2016) explicitly acknowledge prosecutors exploiting legal rules about the type of evidence admissible in drug mandatory minimum cases to secure longer and more certain sentences.

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

III. Data

To estimate the degree of bunching at the 10-year mandatory minimum threshold, I use data on federal cases that include the amount of drugs recorded at sentencing. I then bring in several other datasets from different stages in the criminal justice process to estimate who is responsible for the bunching at 280g.[17]

Figure A1 shows a simplified timeline from arrest to sentencing and describes how the data I use is related to each step. This timeline also acknowledges that selection into/out of the data can occur at

[17] I am not able to link defendants/offenders across these datasets. However, given the nature of the findings and the information available in each dataset, analyzing them independently is sufficient to show where the bunching first occurs and to rule out alternative explanations. Finally, a dataset of defendants/offenders linked from arrest to sentencing does exist, but the codebook for that data suggests it does not include a measure of drug quantity seized at arrest.

each step. As Knox, Lowe, and Mummolo (2019) discuss, bias in selection into the dataset of interest can distort the ultimate measure of bias. My empirical approach takes any bias in selection as given, and assumes this bias does not change sharply in 2010. I show evidence to this effect: drug selling and crack-cocaine usage does not increase after 2010, drug quantities seized do not increase after 2010, and the composition of cocaine offenses in state/local convictions does not change after 2010. Penalties remain high for offenses involving less than 280g, suggesting that there is little reason for selection into federal sentencing to change pre- versus post-2010. Also, Rehavi and Starr (2012) use linked data to show that the probability a case is filed in federal court and the probability a defendant is convicted is the same for black and white defendants (conditional on arrest). Finally, as long as selection into the data is biased in favor of white defendants (i.e. police are more lenient with white defendants or prosecutors are more likely to dismiss white defendant cases), then the estimate of the racial bias in this paper will be an underestimate.

A. United States Sentencing Commission (USSC) Data

To estimate the degree of bunching at or above 280g, I use data provided by the USSC on recorded drug amounts in all federal drug cases sentenced from 1999-2015.[18] I focus on cases that involve a crack-cocaine offense since that is the only drug for which the mandatory minimum threshold changes over time. Approximately 7.8% of offenders in this sample are labeled as white, 10.6% as Hispanic, and 81.6% as black. Table 1 summarizes additional information about age, education, citizenship, and details about the offense, all of which are used as covariates in later analyses (see Appendix D for further details on this dataset and others).

I restrict these data to cases in which the amount of drugs is non-missing and is not recorded as a range. Approximately 20% of cases are excluded for this reason, but the fraction of missing cases for crack-cocaine does not change discontinuously at 2010, though it does increase in 2013 and 2014. Furthermore, in Appendix A, I show that including cases coded as a range only exacerbates the degree of bunching and the racial disparity in bunching. I also remove cases that are flagged for having data issues with the drug quantity variable and cases where the court does not accept or changes the findings of fact. Less than 2% of cases are excluded for these reasons.

Using the cleaned data, I plot two histograms (Figures 1a-b) that zoom in on the density around 280 grams for the years before and after 2010. Prior to 2010, the density around 280g is smooth. After 2010, however, 280g becomes the new mandatory minimum threshold and in that same time, the number of cases at and above 280g spikes.[19] Figures 1c-d display how the fraction of cases recorded as 280-290g

[18] These amounts are derived from pre-sentence reports prepared by a probation officer and in consultation with the defendant, the defendant's counsel, and the prosecuting attorney. 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.

[19] See Figure A3 for a plot of the histogram from 0-500g.

changes over time. This shows even more clearly that the spike in cases at 280-290g coincides exactly with the policy change. These figures also highlight the racial disparity in bunching at the threshold that occurs after 2010.

B. Additional Data

In addition to data on federal sentences from the USSC, I incorporate several other datasets to understand the source of the bunching in drug trafficking cases. I describe these datasets here.

Florida State Inmate Database, 2000-2015.

These data include the year an offender is convicted, a description of the offense, and the offender's race. In Florida, drug offense descriptions typically include the name of the drug involved, and occasionally, the descriptions include a range for the amount of drugs involved (these broad ranges are: 0-28g, 28-200g, 200-400g, and 400+g). Also, Florida does not separately categorize crack versus non-crack cocaine offenses and instead describes all such drug offenses as "cocaine."[20] The fraction of all cocaine cases from 200-400g still exhibits a sharp increase in the USSC federal data, and thus, a mirrored decrease should be detectable using the broad categories in Florida. Summary statistics for these data, the NIBRS drug seizures, and the DEA drug exhibits are reported in Table A1.

National Incident Based Reporting System (NIBRS) Property Segment, 2000-2015.

The FBI collects data from local law enforcement agencies about crime, and many agencies report this data at the incident-level. The incident-level reports make up the data in the NIBRS property segment. These data are submitted voluntarily by agencies and thus, are not representative of national or state-level crime. For this reason, I use a balanced panel of agencies from 2000-2015. Upon receipt, the FBI checks the reports for errors and contacts agencies for corrections if necessary. The property segment of this database includes information about drug seizures and drugs involved in arrests.[21] The offender segment of this database includes information on offender race, sex, and age for all offenders involved in the incident.[22]

DEA System to Retrieve Information from Drug Evidence (STRIDE), 2000-2015.

The STRIDE database contains information about all drug evidence from the DEA and other agencies that was submitted to DEA laboratories for analysis. I obtained the data from a Freedom of Information Act request for all records pertaining to the drug "cocaine" from 2000 to 2015. This information includes

[20]Data from Missouri Department of Corrections indicates that, in Missouri, approximately 80% of state-level cocaine offenses are crack-cocaine offenses.

[21]See Shively (2005) and Bibel (2015) for a discussion of well-known issues with NIBRS data, such as reporting and measurement of hate crimes and sexual assault, differential coverage, and data quality. To the best of my knowledge, there are no known issues with the drug quantity field of the NIBRS property segment.

[22]For tractability, I limit the offender segment to incidents that involve 5 or fewer offenders. This covers 99% of all incidents. Also, the fraction of incidents with 5 or fewer offenders does not meaningfully change after 2010 (99.1% in 2000, 99.1% in 2005, 99.0% in 2010, and 99.3% in 2015). Finally, it is not correlated ($\rho = 0.0001$) with the probability an incident involves 280-290g of crack-cocaine.

the year and month the drugs were acquired, the weight of the drugs in grams, the type of drug (cocaine, cocaine hydrochloride, cocaine base, etc.), drug potency, and the price from undercover purchases.

Executive Office of the US Attorney (EOUSA), Caseload Data, 2000-2017.

The EOUSA releases case-level data on cases (excluding certain redacted cases) processed by the US Attorney's office. These data are derived from information entered into the Legal Information Office Network System (LIONS) case management system. The EOUSA notes that each district may use LIONS differently, and as such, the data should not be used to make cross-district comparisons. The analyses using these data are robust to the inclusion of district fixed effects and various methods of accounting for missingness in the drug quantity data (a data quality issue that varies across districts). The EOUSA data includes a wealth of information about drug cases and other cases, including type of drug, quantity of the drug, an ID for the lead attorney on the case, and an ID for the judge on the case. Summary statistics are reported in Table A2.

National Survey on Drug Use and Health (NSDUH), 2002-2016.

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

Google Search Trends Data on Racial Animus from Stephens-Davidowitz (2014), 2004-2007

To measure racial animus at the state-level, I use data introduced by Stephens-Davidowitz (2014). Stephens-Davidowitz uses Google search data from 2004-2007 (accessed via the Google Trends tool) and measures 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. In fact, it is positively correlated with racial animus as measured by implicit association tests or questions about interracial marriage from the General Social Survey.[23] Even more, it is highly predictive of President Obama's vote share in the 2008 and 2012 US elections (Stephens-Davidowitz 2014). The construction of the measure is covered in much greater detail in Stephens-Davidowitz (2014).

Implicit Association Test (IAT) Data on Racial Animus for Lawyers, 2006-2016.

The IAT data from Project Implicit (Xu et al. 2019) contains the results of implicit association tests for racial bias for over 3 million individuals. The implicit association test for racial bias is designed to test how strongly a person links black people with the concept of "bad" and white people with the concept of

[23] It is also correlated at the Census region level with responses to these questions from respondents with a graduate degree. This suggests it is not solely reflective of racial animus from people with low levels of education. See Figures A4a-i.

"good." This is accomplished by having a person sort words into "good" and "bad" categories, sort people into "black" and "white" categories, and finally, sort both words and people into "black" and "white" categories paired with "good" or "bad" categories. The time it takes to sort into "black/good" relative to "black/bad" and "white/bad" relative to "white/good" is the basis of a person's score. See "Project Implicit" for more detail. Although recent research casts doubt on the validity of the IAT for detecting bias (Oswald et al. 2013), the data has two advantages. First, it can be aggregated to the federal district, a sub-state geography. Second, it can be calculated solely for people reporting an occupation of "Lawyers, Judges, and Related Workers."

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 observed evidence or a difference in the likelihood a case is bunched **conditional** on the observed evidence (i.e. a **conditional racial disparity**). Third, to estimate who causes the bunching at 280g after 2010. And fourth, to explore and test various explanations for the racial disparity in bunching, including discrimination. In this section, I detail methodology for the first three goals. I reserve the discussion of potential discrimination and related tests for Section V.E.

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 any 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 drug quantity distribution above 50g. Section IV.B outlines tests for a conditional racial disparity under this assumption.

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

A. Bunching at 280g and Racial Disparity in Bunching

I define a case as "bunched" at 280g as any case in the narrow range 280-290g (not including 290g). I then compare the fraction of cases from 280-290g in the post-2010 distribution of drug weights to the fraction of cases from 280-290g in the pre-2010 distribution. Specifically, I estimate the following linear probability model:

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

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.[24] After2010_{it} is equal to one if the offender **i** in year **t** is sentenced in 2011-2015 and is equal to zero if the offender is sentenced in 1999-2010. β is the change in an offender's probability of being charged with an amount in the narrow 280-290g range as a result of being sentenced after the threshold amount is increased to 280g. Z_i represents case-level covariates (such as offender education, race, age, conviction state, etc), and $g(t)$ represent time trends. In most specifications, I limit the sample to 0-1000g to remove extreme outliers and exclude Z_i and $g(t)$, however I show that the result is robust to altering this sample range and robust to including numerous controls.

To estimate heterogeneity in bunching by race, I extend the model as follows:

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

Now, β represents the change in a white offender's probability of being charged with 280-290g as a result of being sentenced after the threshold is increased, and δ represents the change for black and Hispanic offenders.[25]

Models (1) and (2) quantify the excess mass at 280-290g by using regression analysis on the case-level microdata and comparing the pre- and post-2010 distributions. This follows work by: Kleven et al. (2011), Behagel and Blau (2012), Sallee and Slemrod (2012), Chetty, Friedman, and Saez (2013), Dwenger et al. (2016), Goncalves and Mello (2018), and Traxler et al. (2018). This approach is also appropriate for the empirical setting. I am primarily interested in estimating the change in the probability a case is charged with 280-290g after 2010 and whether that change in probability differs by race. In addition, some analyses in the paper preclude aggregating the data into bins because they rely on data

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

[25] Combining black and Hispanic offenders into one category, although common in analyses of the criminal justice system, is a crude categorization. Splitting these groups into separate variables yields similar results. There is a larger increase in bunching for black offenders than white offenders and a larger increase for Hispanic offenders than white offenders. The increase in bunching is similar for black and Hispanic offenders. For expositional reasons, I combine these groups throughout the paper. However, it is worth noting that these groups' 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).

that do not include precise drug quantities.[26]

To understand where the excess mass at 280-290g comes from (i.e. where the post-2010 distribution has less mass relative to the pre-2010 distribution), I estimate a series of models similar to the equation (1) that replace the dependent variable with different drug quantity ranges:

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

In these models, β represents the change in an offender's probability of being charged with an amount of drugs between X and Y grams as a result of being sentenced after the threshold is increased. I estimate equation (3) for 0-5g, 5-28g, 28-50g, 50-60g, 60-100g, 100-280g, 280-290g, 290-470g, 470-600g, and 600-1000g. The prosecutor objectives discussed in Section II.B yield specific predictions about many of these ranges–an increase in the 0-5g and 280-290g ranges and a decrease in the 50-60g, 60-100g, and 100-280g ranges.[27] The missing mass analysis addresses a critical question for policy implications: how would offenders who were charged with 280-290g post-2010 have been charged pre-2010? If those offenders would have been charged below 280g, then the bunching at 280-290g post-2010 may represent an effort to increase sentence lengths for some offenders.

B. Racial Disparity Conditional on Observed Drug Behavior

Now, I outline the assumptions necessary to estimate whether the racial disparity in bunching at 280g is due to differences in the underlying distributions of observed evidence or a difference in the likelihood a case is bunched **conditional** on the observed evidence.

1. Institutional Setting

Consider a simplified criminal court setting with drug cases, prosecutor discretion over amount charged, and mandatory minimum sentences. Assume the **seized evidence** s in a case is drawn from a discrete distribution $G_r(.)^t$ that is specific to each race r and time-period t (pre- vs. post-2010). The prosecutor for the case chooses the **amount (in grams) of drugs charged** a, and can charge amounts higher than s by collecting additional evidence $a-s$. Seized evidence s is a noisy measure of true drug trafficking.

[26]In Appendix B, I show that the results in this paper are robust to alternative methods of quantifying bunching above the threshold. One approach, introduced by Saez (2010) and Chetty et al. (2011), constructs a high-order polynomial counterfactual density from the actual bunched density. Kleven (2016), however, notes that this standard bunching estimation is typically used in settings where there is no variation in the kink/notch, and calls this a "minimalist approach" that "may not be compelling in all contexts." Additionally, he argues "more sophisticated alternatives exist that require richer data and/or richer variation." The Fair Sentencing Act in 2010 provides richer variation in this setting. A second alternative approach takes advantage of that variation by aggregating the post-2010 distribution and the scaled pre-2010 distribution into 10g bins and comparing them directly in levels. The results in this paper are robust to both.

[27]In Appendix A, I report the analysis by race for more narrow ranges. Since the ranges involved are much wider than the previous bins, I include a time trend (centered at zero in 2011) and state fixed effects to account for broad differences in drug trafficking over time and across states. In some specifications, I also estimate the "jump" in the probability of being below or above the 280-290g range after 2010. This approach yields similar results, and it is discussed in more detail in Appendix A.

I observe the amount charged a. Publicly available data from the USSC does not report the seized evidence s for each case, and true drug trafficking is unknown to the researcher and the prosecutor. The prosecutor chooses a based on a variety of factors. The first goal of the empirical analysis is to identify racial disparities in a conditional on s (i.e. a **conditional racial disparity**). The second goal (addressed in Section V.E) is to model under what conditions the disparity reflects discrimination by prosecutors and to conduct empirical tests of that model.

In this section, I detail the identifying assumptions necessary to estimate the conditional racial disparity. The set-up closely follows Goncalves and Mello (2018) who use a difference-in-bunching design to estimate police officer bias in speeding tickets. For now, consider the prosecutor's objective a function of tastes (including racial biases), career concerns, the sentence that would be justified under law if true drug trafficking were observed, and costs associated with building the case.

The amount of drugs charged a maps onto a **mandatory minimum sentencing schedule** $l(a)^t$ that differs pre-2010 $t = 0$ and post-2010 $t = 1$.

$$l(a)^t = \begin{cases} 1 & \text{if } a < mm_L^t \\ 5 & \text{if } mm_L^t \leq a < mm_U^t \\ 10 & \text{if } mm_U^t \leq a \end{cases} \tag{4}$$

If a is below the lower threshold for time period t, the defendant is sentenced to 1 year. If a is equal to or above the lower threshold but below the upper threshold, the defendant is sentenced to 5 years. If a is equal to or above the upper threshold, the defendant is sentenced to 10 years. A mandatory minimum does not, by law, require a discontinuous increase in sentence length at the thresholds. In practice, sentences do jump at 50g pre-2010 and 280g post-2010.

Given the seized evidence s (unobserved in the data but observed by the prosecutor) in the case and the defendant's race r (observed in the data), the prosecutor charges a final amount a (observed in the data) that is equal to a mandatory minimum threshold $mm = \{5, 28, 50, 280\}$ (i.e. "bunching" at the threshold) with a **bunching probability** $Pr(a = mm|s, r)^t$ (unobserved in the data). Finally, let defendants be in one of two broad **race** categories: white $r = w$ or black/Hispanic $r = bh$.

2. Defining the Conditional Racial Disparity

Now, I define a racial disparity in the amount charged a conditional on s and outline key equations.

There is a **conditional racial disparity** in bunching at 280g after 2010 if $Pr(a = 280|s, bh)^1 > Pr(a = 280|s, w)^1$. In other words, a conditional racial disparity exists if a black or Hispanic defendant with amount seized s is more likely to be bunched at 280g than a white defendant with the same amount seized s.

I observe the final amount charged, which can be written for the following ranges as:

$$
Pr(a=j|r)^t = \begin{cases} \begin{cases} (a) \quad Pr(s=50|r)^0 + \sum_{k<50} Pr(s=k|r)^0 \times Pr(a=50|s=k,r)^0 & \text{if } j=50 \\ (b) \quad Pr(s=j|r)^0 & \text{if } 50<j \end{cases} & \text{if } t=0 \\ \begin{cases} (c) \quad Pr(s=j|r)^1 \times (1-Pr(a=280|s,r)^1) & \text{if } 50< j<280 \\ (d) \quad Pr(s=280|r)^1 + \sum_{k<280} Pr(s=k|r)^1 \times Pr(a=280|s=k,r)^1 & \text{if } j=280 \\ (e) \quad Pr(s=j|r)^1 & \text{if } 280<j \end{cases} & \text{if } t=1 \end{cases} \tag{5}
$$

Equations **(5.a)** and **(5.b)** express the probability a case is charged with a given amount a prior to 2010. First, the probability a defendant is charged with an amount a equal to 50g is equal to the probability the seized evidence s is 50g plus the likelihood that a case with s under 50g gets moved up to 50g **(eqn. 5.a)**. Second, since there is no sentencing benefit of charging an amount above 50g, the probability a case is charged above 50g **(eqn 5.b)** is equal to the probability s is equal to that amount.

Equations **(5.c)-(5.e)** express the probability a case is charged with a given amount a after 2010. The probability a case is charged with an amount below 280g and above 50g **(eqn. 5.c)** is equal to the probability that s is equal to that amount and that the case does not get moved up to 280g given the amount s. The probability a case is charged with 280g **(eqn. 5.d)** is equal to the probability s is 280g plus the likelihood that a case with s under 280g gets moved up to 280g. As in **(eqn. 5.b)**, the probability a case is charged above 280g **(eqn 5.e)** is equal to the probability that s is equal to that amount. Throughout, I assume that prosecutors don't suppress evidence, i.e. $a \geq s$.[28]

3. Difference-in-Bunching Estimator and the Conditional Racial Disparity

To estimate whether $Pr(a=280|s,r)^1$ differs for black/Hispanic vs. white defendants, I compare the distribution of amounts charged after 2010 to the distribution of amounts charged prior to 2010.

Under the assumption that $Pr(s=k|r)^0 = Pr(s=k|r)^1$–i.e., the probability a case with a defendant of race r has seized evidence $s=k$ does not change pre- vs. post-2010–the **difference-in-bunching coefficients (eqn. 2)** $\delta - \beta$ yields the following:

[28] In reality, it is possible for prosecutors to reduce the drug amount charged or choose not to pursue a drug charge entirely. Introducing this possibility means the disparity in bunching could be due to: (1) a difference in underlying observed drugs, (2) a conditional disparity in bunching, or (3) a conditional disparity in suppressing. The empirical evidence I show is consistent with (2) and (3), both of which are disparities conditional on underlying observed drugs. For that reason, I focus on the simpler case.

$$\delta - \beta = [\sum_{k<280} Pr(s=k|bh) \times Pr(a=280|s=k,bh)^1] - [\sum_{k<280} Pr(s=k|w) \times Pr(a=280|s=k,w)^1] \tag{6}$$

$\delta > 0$ and $\beta > 0$ imply that prosecutors increase a in response to the Fair Sentencing Act, and $\delta - \beta > 0$ implies that they increase a more for black and Hispanic defendants. This alone is of interest–it shows that prosecutors use their discretion to increase sentences in response to the FSA and that the burden of this falls on minority defendants. However, $\delta - \beta > 0$ could be driven by different underlying distributions of seized evidence s (i.e. different $Pr(s = k|r)$) or by disparate treatment conditional on s (i.e. different $Pr(a = 280|s, r)^1$–a conditional racial disparity).

The goal of this section is to outline how to test whether $\delta - \beta > 0$ is due to a conditional racial disparity. I detail two tests. For the first test, $\delta - \beta$ can be rewritten as follows:

$$\begin{aligned}\delta - \beta = & \overbrace{[\sum_{k\leq 50} Pr(s=k|bh) \times Pr(a=280|s=k,bh)^1 - \sum_{k\leq 50} Pr(s=k|w) \times Pr(a=280|s=k,w)^1]}^{H} \\ & + \overbrace{[\sum_{50<k<280} Pr(s=k|bh) \times Pr(a=280|s=k,bh)^1 - \sum_{50<k<280} Pr(s=k|w) \times Pr(a=280|s=k,w)^1]}^{I}\end{aligned} \tag{7}$$

First, I test whether the H term can explain $\delta - \beta > 0$. I observe $Pr(a = 50|r)^0$ and $Pr(a = 50|r)^1$. Equation **(5)** implies that:

$$\begin{aligned}Pr(a=50|bh)^1 - Pr(a=50|bh)^0 = & -[Pr(s=50|bh) \times Pr(a=280|s=50,bh)^1] \\ & -[\sum_{k<50} Pr(s=50|bh) \times Pr(a=50|s=k,bh)^0]\end{aligned} \tag{8}$$

Under the assumption that $Pr(a = 50|s, r)^0 \geq Pr(a = 280|s, r)^1$ for all $s \leq 50$, equation **(8)** is greater than the $\sum_{k\leq 50} Pr(s = k|bh) \times Pr(a = 280|s = k, bh)^1$ term from equation **(7)**. Thus, if the sum of equation **(8)** and $\delta - \beta$ is greater than zero, then the term H cannot explain $\delta - \beta > 0$. In other words, the shift **from** 50g for black and Hispanic offenders is an upper bound for the movement **to** 280g that can be explained by amounts seized at 50g or below. If this shift is not enough to explain the racial disparity in bunching at 280g, then the racial disparity must be due to term I.

Second, I test whether racial differences in $\sum_{50<k<280} Pr(s = k|r)$ from term I can explain $\delta - \beta > 0$. From equation **(5.b)**, $Pr(a = k|r)^0 = Pr(s = k|r)^0 \; \forall \; 280 > k > 50$. Thus, I can test if $\sum_{50<k<280} Pr(s = k|w)^0 = \sum_{50<k<280} Pr(s = k|bh)^0$ by testing if $\sum_{50<k<280} Pr(a = k|w)^0 = \sum_{50<k<280} Pr(a = k|bh)^0$. In other words, if the distributions of pre-2010 charged amounts from 50-280g are approximately equal by race, then the racial disparity in bunching must be due to a racial disparity in the probability a case is bunched at 280g **conditional** on the seized evidence.

Now, I turn to the second test for a conditional racial disparity. The assumptions above also imply:

$$Pr(a = 50 < k < 280|r)^1 = Pr(s = k|r)^1 \times (1 - Pr(a = 280|s = k, r)^1) \tag{9}$$

$$Pr(a = 50 < k < 280|r)^0 = Pr(s = k|r)^0 \tag{10}$$

The difference between equation **(9)** and **(10)** by race can be estimated as follows:

$$(\text{Charged X-Yg})_{it} = \alpha + \delta^X(\text{After2010} \times \text{BlackOrHispanic})_{it} + \gamma\text{After2010}_{it} + \lambda\text{BlackOrHispanic}_i + \epsilon_{it} \tag{11}$$

The coefficient $\delta^X = Pr(a = 280|w, s)^1 - Pr(a = 280|bh, s)^1$. Then, $\delta^X < 0$–i.e., black and Hispanic defendants are more likely to be shifted away from a given amount X after 2010–implies that there is a racial disparity in amount charged a conditional on the underlying evidence seized s.

V. Results

A. Main Results

1. Primary Bunching Estimates and Robustness

Using final sentencing data from the USSC, I estimate the effect of being sentenced after 2010 on whether an offender is sentenced for a drug amount between 280-290g. Column 1 of Table 2 indicates that offenders sentenced after the threshold increases to 280g are more likely to be charged with amounts just above 280g. An offender sentenced after 2010 is 3.5 percentage points more likely to be charged with a drug amount between 280-290g. Column 2 shows that this increase in bunching is driven by black and Hispanic offenders, who are approximately three times as likely to be charged with 280-290g after 2010 compared to white offenders. Figures 1a-d display graphical evidence of bunching at 280-290g and the racial disparity in that bunching.[29]

This result is robust to various sample restrictions (e.g. limiting to post-2006 years); the inclusion of state fixed effects, time trends, state-specific time trends, and offender-level controls (e.g. education, criminal history, age, etc.); clustering standard errors at the state-level; the use of Logit/Probit/Poisson

[29] Figures A5-A7 and B1-B4 present alternative ways to visualize this phenomenon. In particular, Figure A6 shows that the total number of cases at 280-290g increases after 2010.

models instead of a linear probability model; wider definitions of the bunching range (e.g. 280-380g); and the inclusion of cases with weights coded as range. See Tables A3-A7 for these results. I also conduct a simple bounding exercise in Table A8 that accounts for potential substitution into other drug types or selection into the case's drug weight being coded as a range. Table A9 presents a difference-in-differences analysis of bunching using other drug types for which the mandatory minimum threshold did not change. These additional tests confirm the main results. Offenders sentenced after 2010 are more likely to be charged with 280-290g, and this increase is disproportionately large for black and Hispanic offenders.

2. Source of the Excess Mass at 280g

To understand the reason for this bunching at 280g, I analyze other parts of the drug quantity distribution. If the excess mass in 280-290g after 2010 comes from above 290g, bunching may be the result of negotiation between prosecutors and defendants (Bjerk 2017). However, if the excess mass comes from below 280g, it is possible that prosecutors are shading amounts upward to exceed the threshold and secure longer and/or more certain sentences.[30]

In Table 3, I show the change in the probability of being recorded in several different ranges: 0-5g, 5-28g, 28-50g, 50-60g, 60-100g, 100-280g, 290-470g, 470-600g, and 600-1000g. Table 3 shows that the probability a case is recorded in those ranges matches the conceptual discussion in Section II.B.[31] In Figures A7a-i, I plot the share of cases over time in each of these ranges. I estimate the regressions in Table 3 by race in Table A10a. The results are similar but noisier since it requires cutting the already narrow ranges by race. Table A10b and Figures A7j-k shows results by race using broader ranges: 0-280g and 290-1000g. In Table A10c, I re-estimate Table 3 including only years from 2007-2015, and I find similar results.

Summing the coefficients in columns 4-6 of Table 3 implies that the change in probability from 50g-280g can account for 87% of the increase in the 280-290g bin. Is it possible that some offenders charged with 280-290g post-2010 would have been charged below 50g prior to 2010? A fixed cost

[30] To be clear, it is impossible to say with certainty that the "missing mass" in the distribution is where cases in the "excess mass" would be recorded had they been sentenced prior to 2010. This is true for nearly all bunching analyses (panel bunching designs that follow the same unit over time are more convincing in this respect). As is typical in bunching analyses, I assume that the missing mass is indicative of where the "excess" cases would be located in the counterfactual. This is not guaranteed by the research design. Instead, this is another piece of suggestive evidence that the bunching is a result of cases being shifted in a way that is consistent with a simple conceptual model of prosecutor behavior and the empirical evidence of no offender response.

[31] Although it is not clear from these analyses, there is excess mass at 50g (the pre-2010 threshold) even after the threshold changes in 2010. This persistent excess mass at 50g is likely due to round-number bias from offenders, police, or prosecutors. The powder cocaine distribution, which never has a mandatory minimum threshold at 50g, exhibits similar excess mass at 50g. For crack-cocaine, the fraction of cases from 50-60g is about 1.5 times the fraction of cases from 40-50g. For powder cocaine, that ratio is similar–the fraction of cases from 50-60g is about 1.7 times the fraction of cases from 40-50g. While conventional bunching estimation would address the presence of round-number bias by accounting for it in the estimation of the polynomial counterfactual, the difference-in-bunching method accomodates round-number bunching directly because that bunching will be present in both the counterfactual (pre-2010) and actual (post-2010) distributions (Best et al. 2018).

of evidence-gathering could explain this behavior. For example, if an offender is arrested with 10g of physical evidence prior to 2010, it may not be worthwhile to collect evidence to push them from a 5-year sentence to a 10-year sentence. After 2010, however, that same offender would face a 1-year sentence without additional evidence-gathering. Once prosecutors pay the fixed cost to gather evidence, it may then be worthwhile to gather enough evidence to reach the 10-year sentence.

Finally, I examine the degree of bunching in the subset of cases that go to trial. If the bunching is a result of lenient prosecutors rounding down, we should expect less bunching in trial cases where incentives for leniency are muted. However, the degree of bunching and the racial disparity in bunching is only heightened in trial cases (see Column 3 of Table 2). In fact, the only cases with 280-290g that go to trial are those of black and Hispanic offenders. As before, the increased bunching is accompanied by a falling share of cases below 280g ($\beta = -0.109$ and $SE = 0.022$) and a small, rising share of cases above 290g ($\beta = 0.034$ and $SE = 0.019$).[32] This is further evidence that the observed bunching is a result of shading up rather than negotiating down. In Section V.C.3, I show additional evidence from prosecutor case management data that cases bunched at 280g would likely be recorded below 280g in the absence of strategic prosecutor behavior around the mandatory minimum threshold.

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

The results above indicate that there is a racial disparity in bunching at 280g. However, those results alone are not enough to understand why there is a racial disparity in bunching. It could be that there are different underlying distributions of observed drug behavior by race. For example, suppose black and Hispanic defendants are more likely to be arrested with 200g and white defendants are more likely to be arrested with 100g. If defendants with 200g are more likely to be moved to 280g, then a racial disparity will emerge. On the other hand, suppose that among defendants with 200g, black and Hispanic defendants are more likely to be moved to 280g–this would imply there is a disparity in bunching conditional on observed drug amount.

Section IV.B outlines the assumptions and empirical tests necessary to estimate the conditional racial disparity in this setting. I conduct both tests outlined in that section, and both tests suggest that the disparity in bunching is driven by a conditional racial disparity rather than racial differences in the underlying distribution of observed drug amount.

The first test relies on decomposing the potential bunching at 280g. For the first part of that test, I estimate the racial difference in the shift away from the 50-60g range. Table 4 reports this result. Black and Hispanic offenders are less likely to be charged with 50-60g after 2010. However, the decrease in the 50-60g range is not large enough to explain the racial disparity in bunching at 280g. Adding

[32] See Table A10d for missing mass results using trial cases only.

the decrease from 50-60g for black and Hispanic offenders in column (1) to the increase to 280-290g for black and Hispanic offenders in column (2) yields a new bunching coefficient of 0.0293. The new coefficient is still about three times larger than the coefficient for white offenders, and it is statistically different from the coefficient for white offenders at the one percent level (p-value = 0.003).

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 = 0.788). Alternative evidence from drug seizure records confirms black and white offenders are seized with similar amounts (see Table 6a and Figure A8a-b). Since the racial disparity in bunching at 280g cannot be accounted for by racial differences in movement from 50g or by racial differences in the distribution from 60-280g, this implies the disparity is a conditional racial disparity.

The second test for a conditional racial disparity in bunching relies on estimating racial differences in movement away from other narrow ranges. Figure 2b plots the coefficients from equation **(11)** 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. This implies that at those amounts, black and Hispanic offenders are more likely to be bunched at 280g than white offenders. Again, this implies there is a conditional racial disparity in bunching at 280g.

B. Sentencing Consequences

In order to understand the policy implications of this bunching, I estimate the sentencing consequences of crossing the mandatory minimum threshold. Since mandatory minimum sentencing only gives guidelines about minimum sentencing, it is possible that being above the amount has no effect on actual sentencing.[33] I investigate this by estimating the following:

$$Sentence_i = \alpha + \beta_1 Above280_i + \beta_2 Amount_i + \beta_3 (Above280 \times Amount)_i + \epsilon_i \tag{12}$$

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. For the main results, I focus on cases sentenced after 2010. In Table A11, I estimate similar regressions using the pre-2010 data. 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

[33] In other words, judges could choose to treat defendants with 270g the same as defendants with 280g and apply the mandatory minimum sentence of 10 years to both.

mandatory minimum threshold after 2010. The bunching above 280g suggests this assumption may be violated. As such, I also estimate (12) 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 11.2 years. Figure 3a 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 3b shows that there is no discontinuity in predicted sentence, where sentence is predicted from a model using pre-2010 cases and several offender characteristics. Figure 3c 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.[34,35] See Figure A9 for robustness to bandwidths from 10g to 250g.

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, the results in Section V.A.2 suggest 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. Using that value as the counterfactual sentence implies a sentencing consequence of 3.3 years.

C. Potential Mechanisms

The four mechanisms I evaluate are: (1) offender responses to the FSA, (2) a shifting composition of cases between state and federal court, (3) law enforcement discretion, and (4) prosecutorial discretion. For these analyses, I present visual evidence as well as a formal analysis of the microdata showing the main bunching results for each mechanism in Table 5 and Tables 6a-b. Ultimately, I find bunching at 280g in prosecutor case management files from the EOUSA but not at an earlier stage. This implies that prosecutors are responsible for the excess mass at 280g in final sentences. In Section V.C.4, I discuss several additional empirical tests that also suggest prosecutors are responsible for the bunching of cases at and above 280g.

1. Offender Behavior

If black and Hispanic offenders respond differently than white offenders to the Fair Sentencing Act, a racial disparity in bunching at 280g may reflect prosecutors' reactions to those different responses rather

[34] This is possible because although offenders are negatively selected (in terms of sentence) on some characteristics, like race, they are positively selected on others, like criminal history score.

[35] These estimates indicate that there is a sentencing penalty for crossing the mandatory minimum threshold (both before and after 2010), not that sentences were longer after 2010 or that the sentencing penalty of triggering the mandatory minimum was higher after 2010.

than racial discrimination. In Table 6a, I show that black and Hispanic offenders are not arrested with more drugs following the Fair Sentencing Act, but instead, are holding slightly smaller amounts when arrested after 2010 (after controlling for state fixed effects, sex, and age).[36] In Table 6b, I show that black and Hispanic respondents to the NSDUH are not more likely to report having ever used crack, selling drugs in the past 12 months, or having used crack and selling drugs after 2010. This implies that the racial disparity in bunching cannot be attributed to differential responses in drug-carrying by race.[37]

2. Shifting of Cases Between State and Federal Courts

Drug Convictions in Florida Courts

The USSC data covers the universe of federal drug cases, but it is possible that the type of cases prosecuted in federal court versus state court changes after 2010. Cases can be prosecuted federally for many reasons (see Appendix D for a discussion of the reasons a case can enter federal court). State and local authorities could send more of their high weight, 280g cases to federal court after 2010. Similarly, federal prosecutors could pull more of these types of cases from state and local courts after 2010.

To test this possibility, I use state-level data on cocaine offense convictions from Florida.[38] Florida classifies drug offenses using broad ranges: 0-28g, 28-200g, 200-400g, and 400+g. The USSC data show a sharp 3.6 percentage point increase in cases with 200-400g convicted in a Florida federal district after 2010 (see Table 5, column 7 and Figure A11a). If the bunching in federal cases is due to state and local authorities sending more 280g cases to federal prosecutors, then there should be a mirrored decrease in the fraction of state-level cases in Florida with 200-400g. Even more, the decrease should be especially pronounced for black and Hispanic offenders.

I do not find a decrease in state convictions for 200-400g in general or by race. Figure 4a plots the share of all cocaine cases in Florida that are for offenses with 200-400g of cocaine by race. Columns 1 and 2 of Table 5 confirm this. The probability a state-level drug conviction is in the 200-400g range in Florida does not meaningfully change after 2010. This implies that shifting from state and local courts to federal courts cannot explain the sharp rise of cases at 280g in federal sentencing.[39] In Table A12

[36] Likewise, I find no evidence of a response in the DEA STRIDE data on drug amounts or drug prices (see Figure A10).

[37] Other papers also find that offenders do not respond or respond only modestly to a change in punishments/sanctions. For example, Lee and McCrary (2017) finds that offenders do not discontinuously decrease offending at age 18, despite a discontinuous increase in the probability of a harsh sentence at that age.

[38] In Appendix A, I show similar results for North Carolina. I do not include NC in the main analysis because many of its drug convictions do not include any information about drug type involved.

[39] Since there are many more cases convicted at the state-level versus federal-level, it is possible that a minor, undetectable shift in Florida would be detectable at the Federal-level. This is not the case for the 200-400g range. First, the state-federal disparity in number of cases is due to states prosecuting more minor possession cases than the federal courts. There are 150 crack or powder cocaine cases in the 200-400g range convicted in federal court districts located in Florida after 2010. There are only 200 cases in this range convicted in Florida state courts after 2010. Re-coding 150 of the 200 Florida cases as if they were not in the 200-400g range does yield a detectable effect. Similarly, re-coding 150 cases not in the 200-400g range as if they were in the 200-400g range also yields a detectable effect. This simple simulation implies that a shift of cases from

and Figure A11b-c, I show these results are robust to alternative sample restrictions and to using similar data from North Carolina.

Bunching by Law Enforcement Agency Sending Case to EOUSA

The EOUSA prosecutor case management files (which I analyze in more detail below) include a field that indicates the law enforcement agency that sends the case to the EOUSA. If the bunching at 280g is caused by a shift from state courts to federal courts, then bunching should only be present in cases with state law enforcement involved. In Figure A11d, I plot the fraction of cases with 280-290g over time by the type of agency involved. I find that bunching at 280g is present in cases with state law enforcement involvement and in cases that are sent from Federal agencies (see Table A13 for a formal test). This is further evidence that the bunching at 280g after 2010 is not the result of state to federal case shifting.

3. Law Enforcement Discretion

NIBRS, Local Law Enforcement Drug Seizures

Using a balanced panel of agencies in the NIBRS data on drug crime, I examine the distribution of drug seizure quantities. If local law enforcement is the source of bunching, I should observe an increase in bunching at 280-290g after 2010. Figure 4b plots the fraction of drug seizures with 280-290g over time and does not show an increase in drug seizures with 280-290g after 2010, in general or by race. These results are also shown in Columns 3 and 4 of Table 5.[40] In addition, only 5 incidents total are reported with 280-290g in the NIBRS after 2010. This suggests that discretion in local law enforcement and drug sting tactics cannot explain the bunching in drug amounts after 2010.

DEA STRIDE, Federal Law Enforcement Drug Seizures

I also test for bunching in drug quantities from the DEA's STRIDE database.[41] This data includes exhibits sent to DEA laboratories from both federal and local law enforcement agencies. Figure 4c plots the share of cocaine exhibits with weights from 280-290g from 2000-2015. There is no increase in exhibits with 280-290g after 2010. Again, Table 5 also shows this result. In fact, there are less than 20

Florida to the federal system would be detectable in the state data. A related concern is that the large number of cases in urban counties may mask shifting in rural counties. I split the analysis by counties with greater than 5000 cocaine convictions from 2000-2015 and counties with less than 5000 cocaine convictions from 2000-2015. I do not find substantial shifting for either group. For small counties (those with less than 5000 cocaine convictions), I find a decrease in cases with 200-400g of about 0.1 percentage points. For large counties (those with more than 5000 cocaine convictions), I find no change in cases with 200-400g (less than 0.02 percentage points).

[40]This result is robust to using only states that have full coverage by 2012 (i.e. states in which all agencies are participating in NIBRS) and 90-100% coverage from at least 2008-2015 (DOJ 2012). See Table A14 and Figure A12.

[41]The analysis in this section uses unvalidated DEA data, and I claim authorship and responsibility for all inferences and conclusions that I draw from this information.

total cocaine exhibits in the DEA data with 280-290g after 2010. This further suggests that local and federal law enforcement are not responsible for the observed bunching at 280g after 2010.

4. Prosecutorial Discretion

Bunching in Prosecutor Case Management Files

The EOUSA provides case-level data extracted from their internal case management system. Using this data, I test for bunching in the quantity of drugs recorded in the case management system. Figure 4d shows that there is a sharp increase in the fraction of cases recorded with 280-290g after 2010.[42] Since I find no evidence of bunching in data from earlier stages, this suggests that the bunching occurs once the case is in the hands of the prosecutor.

Table 5 indicates that the fraction of cases in 280-290g increases by 7.7 percentage points after 2010. This is twice the increase I find in the final sentencing data. This difference is likely driven by missing values in the EOUSA files. Re-coding each missing value as though it were not in the 280-290g range (i.e. equal to zero) yields an increase of about 3.5 percentage points after 2010, which is consistent with estimates from the sentencing data. The main results below are robust to missing value re-coding.[43]

I also examine bunching at 280g for cases received by the EOUSA before the Fair Sentencing Act is signed into law. These cases are less likely to be influenced by offender or police responses to the FSA. For cases that are received by the EOUSA 60 days before the FSA but sentenced after the FSA, 2.7% are bunched at 280-290g. For cases that are received by the EOUSA 60 days before the FSA and sentenced before the FSA, 0.4% are bunched at 280-290g. The timing of bunching in these cases further suggests the increase in bunching at 280g is due to prosecutor decisions.[44]

The EOUSA data do not contain a field for race of the defendant. I can impute race for cases from the EOUSA data that contain a sentence month and year (not all cases received are sentenced) by using the racial composition of sentencing in each year-month from the USSC sentencing data. As before, I find an increase in 280-290g cases after 2010 and a particularly large increase in months with more black and Hispanic offenders sentenced (see Table A15). In Table A15, I also show that the disproportionate bunching for black and Hispanic offenders (using imputed race) is robust to including prosecutor fixed effects.

[42] See Figure A13a for a plot of bunching at 280g by the month the case is received.

[43] See Table A15 and Figure A13b-c for the main bunching results after re-coding the 280-290g dummy variable as equal to zero when the drug weight is missing.

[44] Figure A13d plots bunching by year sentenced for cases received before the FSA.

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 specific circumstances of each case, I cannot pinpoint "bad behavior" from any individual attorney. However, by estimating bunching separately for each attorney, I can calculate the fraction of prosecutors responsible for the observed bunching. Also, I can compare the distribution of cases for bunching and non-bunching attorneys to further understand where the excess mass at 280-290g is coming from (i.e. where there is relatively less mass in the bunching attorney distribution compared to the non-bunching attorney distribution).

Prior to 2010, approximately 0.4% of all cases with a drug quantity less than 1000g were recorded as having 280-290g. I use this statistic as a benchmark to detect attorneys who bunch after 2010.[45] For each attorney, I calculate the percentage of their cases with 280-290g of drugs 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 cases after 2010. Results are similar when using lead attorneys with 5 or more drug cases after 2010 or with 15 or more drug cases after 2010.

The majority of these attorneys exhibit no bunching.[46] In other words, their fraction of cases with 280-290g post-2010 is at or below the pre-2010 average. Approximately 30.4% of prosecutors, however, do have a higher than normal percentage of cases with 280-290g after 2010. Drawing 50 samples (stratified on lead attorney ID and with replacement) from the data and re-calculating the fraction of bunching attorneys in each sample yields a standard error of 0.024. This implies a 90% confidence interval on the estimate of about 26.4-34.3%. Over 50% of these attorneys have two or more cases at 280-290g and over 25% have three or more cases at 280-290g.[47] The fraction of bunching attorneys is also significantly different at the one percent level from the fraction calculated by randomly re-assigning cases to prosecutors (see Figure A15).

In Figure A16, I map the number of bunching attorneys in each state (using attorneys with 5 or more drug cases post-2010 to increase the set of states that have eligible attorneys).[48] The attorney-level bunching cannot be accounted for by district fixed effects. The within-district standard deviation in the

[45] I can also use the district-level pre-2010 average to account for district fixed effects in cases at 280-290g. Even more, I can use each attorney's pre-2010 behavior as their own benchmark to detect bunching post-2010. Both approaches yield similar results.

[46] Figure A14 plots a histogram of the resulting measure for the 128 attorneys who served as lead attorney on at least 10 drug cases after 2010.

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

[48] The number of bunching attorneys in a state is positively correlated with racial animus in that state (see Table A16).

280-290g bunching metric is 0.13, the between-district standard deviation is similar at 0.12, and district fixed effects only explain about 6% of the variance in the attorney-level bunching metric.

Further Evidence on Source of Excess Mass at 280g

In Table 7, I estimate the likelihood a case is charged below 280g, with 280-290g, or above 290g for the bunching versus the non-bunching attorneys. This echoes the approach that Goncalves and Mello (2018) use to formally estimate bunching in speeding tickets in Florida.[49] 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 and (2) attorneys who have an above-average share of cases with 50-60g pre-2010. Definition (2) provides a classification of bunching attorneys that is not mechanically related to the fraction of cases in the 280-290g range.[50]

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

Additional Evidence on Prosecutor-level Bunching

Next, I identify attorneys who switch from one federal district to another federal district, and, using the two definitions above, I test whether bunching is persistent across districts. Definition (2) 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 A20 shows these results. I find that an attorney who bunches at the 10-year mandatory minimum 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. In other words, bunching at the 10-year mandatory minimum threshold is a behavior that persists across districts, suggesting that bunching is related to a characteristic of the prosecutor and not another actor in the district (e.g. police, judge, or defense attorney).[51]

In Figure A17, I examine how other prosecutors in a district change their bunching behavior when

[49]They compare lenient police officers to non-lenient police officers.

[50]In Appendix A, I show that the results using definition (1) are robust to categorizing the prosecutor for defendant **i** as a bunching or non-bunching attorney leaving out defendant **i** from the determination, and that all results are robust to bootstrapping the standard errors to adjust for error in the bunching classification. I also show that these results are robust to using attorneys with 15+ cases or 5+ cases. See Tables A17-A19.

[51]Recall, Table A13 shows that the increase in bunching at 280-290g is similar for most police agencies sending cases. This also suggests that the variation in bunching at the prosecutor level is due to prosecutor choices and not choices made by investigators.

a bunching prosecutor enters. I find that that when a bunching attorney switches into a new district, all other attorneys in that district begin bunching more. To conduct this test, I classify bunching attorneys using data from 1994-1999 and definition (2). I then identify the districts that those attorneys move into, and I study the attorneys in that district after the first bunching attorney moves in post-1999. This means earlier years are over-represented. I show that bunching increases in a district once a bunching attorney enters, but that it does not decrease once the bunching attorney leaves. This is 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. Figure A18 shows that bunching at the 10-year mandatory minimum increased by 60% from 1988-90 to 2010, which is consistent with the practice of bunching being learned over time. The figure notes in Appendix A contain a more detailed discussion of these results.

Finally, in Table A21, I show that attorneys who bunch at 280-290g post-2010 also have more cases bunched at 28-29g (the five-year mandatory minimum) post-2010 and more cases bunched at 50-60g pre-2010 (the pre-2010 ten-year mandatory minimum). Likewise, attorneys who bunch at 50-60g pre-2010 also have more cases bunched at 28-29g post-2010 and 280-290g post-2010. One concern about the estimation of prosecutor-level bunching is that the variation across prosecutors could be due to noise alone, especially since I only require prosecutors to have 10 or more cases after 2010. These results that show prosecutor-level 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.

While it may be surprising that prosecutors could induce this bunching, recall that this ability is explicitly written into federal sentencing guidelines. One tool prosecutors can use to increase the weight used at sentencing is tying the defendant to a larger drug conspiracy. Cases with 280-290g after 2010 are more likely to have a lead charge of "drug conspiracy" than cases with 290g-1000g (see Table A22). Prior to June 2013, the evidence about relevant quantities did not need to satisfy the "beyond a reasonable doubt" evidentiary standard, because the "principles and limits of sentencing accountability under this guideline are not always the same as the principles and limits of criminal liability" (USSC, 2015). A Supreme Court decision in June 2013 changed the evidentiary standard, and I evaluate that change below.

D. 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 presiding 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." I estimate how prosecutors reacted to this decision by comparing the change in bunching around June 17, 2013 to the change around June 17th in other years after 2010. If prosecutors are inflating drug amounts to levels that could not be supported at trial, then there will be a decrease in bunching for cases received after the Supreme Court decision.

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

$$\begin{aligned}(\text{Recorded}\,280-290g)_{it} = {} & \alpha_0 + \beta_1 \text{AfterJune17}_{it} + \beta_2 \text{DaysFrom}_{it} + \beta_3(\text{After}\times\text{DaysFrom})_{it} \\ & + \delta_1(\text{AfterJune17}\times\text{Year2013})_{it} + \delta_2(\text{DaysFrom}\times\text{Year2013})_{it} \\ & + \delta_3(\text{After}\times\text{DaysFrom}\times\text{Year2013})_{it} + D_{it} + \epsilon_{it}\end{aligned} \tag{13}$$

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 is equal to zero if case **i** is received before June 17th of year **t** but after January 1st of year **t**. DaysFrom_{it} is the number of days from June 17th that case **i** is received, and Year2013_{it} is equal to one if case **i** is received in 2013 and is equal to zero if it is received in 2011-2012 or 2014-2016.[52] D_{it} represents day-of-week fixed effects. The coefficient β_1 is the average discontinuity in the fraction of cases with 280-290g after June 17 from 2011-2016. The coefficient δ_1 is the discontinuity that is specific to June 17, 2013–the date of the *Alleyne* decision.[53,54]

Column 2 of Table 8 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

[52] I do not include 2017 in these analyses since the data do not include the full year.

[53] In response to *Alleyne*, Attorney General Eric Holder released a memo in August 2013 instructing US attorneys to decline to charge quantities necessary to trigger the mandatory minimum in cases with low-level and non-violent offenders who have little 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 much larger). Also, using updated EOUSA data, I find that there is no change in bunching after May 12, 2017, the day Attorney General Jeff Sessions rescinded the August 2013 Holder memo.

[54] I do not conduct the traditional RD identifying assumption tests in this section. For one, the EOUSA data contain very few case-level covariates. Even more, the resulting discontinuity, whether it arises from prosecutors rushing to try cases before the Supreme Court decision or solely from prosecutors changing their behavior immediately after the decision, reveals that prosecutors were submitting evidence to judges that they believed would not hold up if submitted to a jury. That said, the density of cases is displayed in Figure A19a.

indicates that, on average, there is approximately no change in bunching after each June 17th from 2011-2016.[55] The next coefficient, labeled "After June 17, 2013", shows the change in bunching that is specific to June 17, 2013. I find that bunching changes discontinuously only after June 17, 2013. In fact, the fraction of cases recorded with 280-290g drops by about 15 percentage points after the ruling in Alleyne. This is also the case for the 120-day and 60-day bandwidth, although as I narrow the bandwidth, I lose precision.[56] Table A23 shows that the decrease in bunching after Alleyne is robust to imputing missing values as zero. Figure A20 shows robustness to additional bandwidth choices and choice of polynomial.

Figure 5 illustrates the large discontinuity in the fraction of cases with 280-290g around June 17, 2013. 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.

E. Discrimination and Alternative Explanations

Now, I introduce a simple model of prosecutor objectives to discuss potential explanations for the racial disparity in bunching at 280g and to motivate empirical tests of those explanations.

1. Model of Prosecutor Objectives

First, I detail the prosecutor's decision problem, which determines the probability $Pr(a = mm|s, r)^t$ that a case with a given amount seized s and defendant race r is charged with an amount a that is equal to the mandatory minimum threshold $mm = \{5, 28, 50, 280\}$. Although I do not estimate any of the parameters in the following model directly, I use it to illustrate channels through which $Pr(a = mm|s, r)^t$ may differ by race and to discuss suggestive empirical tests of those various channels.[57]

The prosecutor for the case chooses the **amount (in grams) of drugs charged** a, and can charge amounts higher than **seized evidence** s by collecting additional evidence $a - s$. Seized evidence s is a noisy measure of **true drug trafficking** d, which is unobservable to the prosecutor. For a given case, prosecutor i chooses the amount of drugs charged a to solve the following problem:

[55] The coefficient on AfterJune17 for 2013 is at least twice as large as the next largest all other years from 1999-2016 (when estimating the non-2013 years separately instead of pooling). See Figure A19b.

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

[57] Note, I write down a static model below, but it can incorporate reputational benefits or reputational costs associated with bunching. The data are not amenable to testing dynamics at the prosecutor-level. I focus on the static problem because it has clear connections to empirical tests I can conduct.

$$\max_a \pi(l(a)^t) - \gamma(r,x) \times c_g(a-s) - c_d(|l(a)^t - (l^*(s,r,x) + \phi_i(r,x))|) \tag{14}$$

The function $\pi(.)$ represents the **career benefits** a prosecutor gets from securing a longer sentence. There are also costs to the prosecutor associated with increasing a, such as the **cost of gathering the additional evidence** $c_g(a-s)$ to build the case. This cost $c_g(a-s)$ is increasing in $a-s$.[58] This cost is determined by other actors the prosecutor must face in the process of working a case. Judges, defense attorneys, juries, witnesses, or other actors in the criminal justice system who are racially biased may present fewer obstacles to entering the additional evidence $a-s$ for cases involving black and Hispanic defendants. Also, if defendants of one race procure better defense counsel, that counsel may make it more difficult for the prosecutor to use additional evidence $a-s$. These **cost differences by race** (and other defendant characteristics) are captured in $\gamma(.)$.

The prosecutor also faces a **psychic cost of deviating** $c_d(.)$ from **the sentence that would be justified by law if true drug trafficking were observed** $l^*(d)$. Since true drug trafficking d is unobservable, prosecutors form an expectation of d by solving a signal extraction problem given the seized evidence s, defendant race r, and other characteristics x. This yields $l^*(s,r,x)$.[59]

Finally, a **prosecutor specific taste parameter** $\phi_i(r,x)$ is added to the sentence $l^*(s,r,x)$, reflecting the prosecutor's animus for defendants based on race r or other characteristics x. Assume that only ϕ_i varies at the prosecutor level.

Writing down the prosecutor's objective function makes explicit the various channels that could cause a conditional racial disparity in the probability a defendant is bunched at 280g. First, the disparity could be due to taste-based racial discrimination: $\phi_i(bh,x) > \phi_i(w,x)$. Second, it could be due to statistical discrimination: $l^*(s,bh,x) > l^*(s,w,x)$. Third, it could be due to racial differences in the cost (to the prosecutor) of building a case: $\gamma(bh,x) < \gamma(w,x)$. All three of the channels could also be related to other characteristics x that are correlated with race r rather than race itself.

2. Empirical Tests of Discrimination and Other Explanation

Other Offender Characteristics

First, I test the explanation that the racial disparity in bunching at 280g is driven by a characteristic correlated with race. To do this, I estimate how bunching differs by various observable offender characteristics. Specifically, I estimate equation (2) fully interacted with binary variables for the following offender characteristics: college education or more, male, above the median age for offenders, offense involves a weapon, above the median criminal history score, above the median number of other current

[58] Again, I assume that prosecutors don't suppress evidence and thus, $a \geq s$.
[59] I model the signal extraction problem in Appendix C.

offenses, and convicted in a state with an above median fraction of black or Hispanic cases pre-2010.

This partially addresses concerns that white and black and Hispanic offender's 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 still exists within education categories, then this suggests that the racial disparity is driven by attitudes about race. In Table 9, I show that the racial disparity in bunching exists even within all of these observably similar groups.

The observable characteristics from the USSC data are only a subset of what the prosecutor observes about a defendant. One concern is that black and Hispanic drug offenders may be more likely to operate in drug organizations or gangs, and that prosecutors may charge offenders from gangs with higher amounts for various reasons. 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 than white federal drug offenders. Also, they are less likely to report income from illegal activities prior to arrest.[60] Also, 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 9, column 8). As in the SIFCF data, white offenders are also more likely to face a conspiracy charge. This further suggests that differences in gang participation by race do not explain the racial disparity in bunching at 280g.

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 racial difference in defense counsel could explain the racial disparity in bunching. The data do not include the offender's type of defense counsel in all years. This information is available for 1999-2002, but in those years, black, Hispanic, and white crack-cocaine offenders are equally likely to be represented by private counsel.[61] The 2004 Survey of Inmates in Federal Correctional Facilities also indicates that private counsel retention is the same by race. Using data from the 1999-2002 USSC files, I construct each district's private counsel retention rate and tag districts as below or above median private counsel retention. I find that bunching and the racial disparity in bunching is similar in places

[60] The SIFCF is a nationally representative survey of inmates in federal prisons. Over 3,000 inmates from 39 federal prisons were interviewed for the 2004 survey. The interviews were conducted by the US Census Bureau on behalf of the Bureau of Justice Statistics. At the beginning of the interview, inmates are told their answers are confidential and that their responses cannot be released to the prison or to anyone else in a way that would identify them. These data contain information on whether the offender was involved in a drug organization/gang. Although the statistics are based on self-reports, it does not appear black and Hispanic offenders report differently than white offenders on other sensitive questions, such as whether police used force during their arrest or whether they have had thoughts of revenge.

[61] 21.0% of white offenders, 22.7% of black offenders, and 21.7% of Hispanic offenders retain private counsel from 1999-2002.

with low and high private counsel retention (see Table A24).

Next, I consider whether the racial disparity in bunching can be attributed to judge bias. I am able to match approximately half of the cases in the EOUSA files to a judge race and political party. For these cases, I do not find any evidence that judge race or political party influences the probability a case is bunched at 280g (see Table A25).[62] Also, unlike prosecutors, judges with a high share of cases at 280g post-2010 are not any more likely to have cases at 28g post-2010 or at 50g pre-2010 (see Table A26).

I also test whether district-level differences in costs of gathering evidence are related to bunching at 280g. I find that the increase in bunching at 280g is similar in districts with a low and high fractions of cases declined due to "weak evidence" or "lack of resources" (see Table A23).[63] This suggests that costs of developing evidence are not related to the rise in bunching at 280g.

Taste-based vs. Statistical Discrimination

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

One potential explanation of these results is that some prosecutors have biased tastes against black and Hispanic drug offenders and believe they should be punished more harshly than white drug offenders. To explore the taste-based discrimination mechanism, I use a state-level measure of racial animus constructed by Stephens-Davidowitz (2014) based on intensity of Google searches including racial slurs in each state. I match this measure to the USSC Sentencing 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 they serve in, and assistant US attorneys have a choice over where to apply.[64]

[62] I have also examined heterogeneity in bunching by race of the head US attorney in the district and the racial composition of prosecutors, judges, defenders, and probation officers in the district. I do not find robust results on these margins.

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

[64] Recall that Alleyne v. US made the jury more important in mandatory minimum cases after 2013. This change led to stricter evidentiary standards for mandatory minimum cases (beyond a reasonable doubt versus preponderance of evidence). However, if juries are, on average, more racially biased than judges, then the effect of Alleyne v. US may be buffered by the increased racial bias of juries. I find that the fraction of cases at 280-290g in low racial animus states (below median) fell by 40% from 2011-2012 to 2014-2017. In high racial animus states (above median), the fraction of cases at 280-290g fell by 20%. This is suggestive evidence that Alleyne was, in fact, less effective in states with high racial animus. However, in all states, the

Again, I estimate equation (2) fully interacted with a dummy variable for high racial animus states that is 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 racial animus is correlated with some state-level preference for harsh sentencing, then I should find an effect for both white and black and Hispanic offenders. However, if the effect is driven by racist beliefs about black and Hispanic offenders, then it should only be present for those groups.

I find that in states with a higher level of racial animus, bunching at 280-290g is more prevalent specifically for black and Hispanic offenders.[65,66] These results are in Tables 9-10. Column 8 of Table 9 shows that in states with high levels of racial animus, black and Hispanic offenders are substantially more likely to be charged with an amount at or slightly above the mandatory minimum threshold.

Table 10 explores the robustness of this result. Columns 1-4 introduce individual and district-level controls interacted with the after 2010 by race dummy variables, and the relationship between animus and bunching is unchanged. Columns 5 and 6 estimate the relationship between bunching and the continuous measure of state-level animus from Google Trends. The coefficient in column 5 is not statistically significant (p-value = 0.2), but the magnitude is much larger than the coefficient for white offenders. Also, based on that coefficient, white and black and Hispanic offenders at low-levels of animus are not statistically different from each other, but they are statistically different at higher levels of animus. Column 6 re-estimates column 5 after eliminating outliers in the animus measure (states with animus below the 1st percentile or above the 99th percentile).

In column 7 of Table 10, I introduce a district-level of racial animus by aggregating implicit association test scores for people reporting an occupation of "lawyers, judges, and related workers." Since many states contain multiple federal districts, I include state fixed effects interacted with after 2010 by race dummy variables. The estimate, then, is identified from within state variation in the IAT animus measure. I find the average IAT score of lawyers in a federal district is correlated with higher bunching for black and Hispanic offenders (p-value = 0.14).

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 and above that point

increase in evidentiary standards led to a net decrease in cases at 280-290g.

[65] The racial animus measure was developed to measure animus against black people. I assume that this is correlated with animus for Hispanic people, so I focus on the pooled results. However, the estimates are similar if I exclude black offenders or Hispanic offenders.

[66] Specifically, I split states by above/below the median racial animus. States above the median racial animus measure are: AL, AR, CT, DE, FL, GA, IL, IN, KY, LA, MD, MI, MO, MS, NC, NJ, NV, NY, OH, OK, PA, RI, SC, TN, and WV. States below the median racial animus measure are: AK, AZ, CA, CO, HI, IA, ID, KS, MA, ME, MN, MT, ND, NE, NH, NM, OR, SD, TX, UT, VA, VT, WA, WI, and WY.

where the mandatory minimum sentence increases, and that bunching is disproportionately larger 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 observed drug amounts.

Since the bunching only appears in prosecutor case management data and the final sentencing data but not in data on state-level convictions or drug seizures, it is likely a result of prosecutorial discretion. Several additional tests confirm this. In fact, just 20-30% of attorneys account for 100% of the bunching observed in the case management data. In addition, bunching becomes less prevalent among prosecutors following a Supreme Court decision that requires stricter evidentiary standards for drug quantity evidence. This, in addition to numerous other tests discussed above, suggests that prosecutors are shading drug amounts upward to induce longer sentences.

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

Finally, the bunching in drug weights and the racial disparity in bunching has meaningful implications for the racial sentencing gap. Depending on the counterfactual sentence imputed for the affected offenders, bunching at 280g can account for 2-7 percent of the racial disparity in crack-cocaine sentences. A highly conservative estimate suggests that being bunched at 280g adds 1-2 years to an offender's sentence. Multiple estimates suggest the cost of incarceration (combining direct care costs and the cost of lost current and future wages for the offender) is approximately $60,000 per person per year (Donohue 2009; Mueller-Smith 2015). I find 3.6% of black and Hispanic crack-cocaine offenders are bunched at 280g after 2010 versus 1.2% of white crack-cocaine offenders. Assuming 3.6% and 1.2% of all drug cases from 1999-2015 were subject to similar discretion by race implies total costs of 1.3 billion dollars for black and Hispanic offenders versus 148 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 2,500 years sentenced for white offenders.

References

Agan, Amanda, Matthew Freedman, and Emily Owens. 2018. "Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense." NBER working paper.

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

Antonovics, Kate L. and Brian 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, Shamena and Hanming Fang. 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 96 (1): 127–151.

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

Arnold, David, Will Dobbie, and Crystal S. Yang. 2017. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics*.

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

Baker, Scott and Claudio Mezzetti. 2011. "Prosecutorial Resources, Plea Bargaining, and the Decision to Go to Trial." *The Journal of Law, Economics, and Organization* 17 (1): 149-167.

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

Barkow, Rachel E. 2008. "Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law." *Stanford Law Rev*iew 61: 869.

Becker, Gary S. 1957. *The Economics of Discrimination: An Economic View of Racial Discrimination*. University of Chicago.

Behaghel, Luc, and David M Blau. 2012. "Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age." *American Economic Journal: Economic Policy* 4 (4): 41–67.

Berger v. United States, 295 U.S. 78. 1935.

Best, Michael, James Cloyne, Ethan Ilzetzki, and Henrik Jacobsen Kleven. 2018. "Interest Rates, Debt and Intertemporal Allocation: Evidence from Notched Mortgage Contracts in the United Kingdom." Working paper.

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

Bibel, Daniel. 2015. "Considerations and Cautions Regarding NIBRS Data: A View From the Field." *Justice Research and Policy* 16 (2): 185-194.

Bjerk, David. 2005. "Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion under Mandatory Minimum Sentencing." *The Journal of Law and Economics* 48 (2): 591–625.

Bjerk, David. 2007. "Guilt Shall Not Escape or Innocence Suffer? The Limits of Plea Bargaining When Defendant Guilt Is Uncertain." *American Law and Economics Review* 9 (2): 305–329.

Bjerk, David. 2017. "Mandatory Minimums and the Sentencing of Federal Drug Crimes." *Journal of Legal Studies* 46 (1): 93-128

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

Carr, Jillian and William McClain. 2018. "Evidence On Court System Bias From Strategic Judge Assignment." Working paper.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics* 126 (2): 749–804.

Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103 (7): 2683–2721.

Dee, Thomas S., Will Dobbie, Brian A. Jacob, and Jonah Rockoff. 2017. "The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations." Forthcoming at *American Economic Journal: Applied Economics*.

Diamond, Rebecca and Petra Persson. 2017. "The Long-Term Consequences of Teacher Discretion in Grading of High-Stakes Tests." Working paper.

Donohue, John 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*, edited by Steven Raphael and Michael A. Stoll, 269–342. New York: Russell Sage Foundation

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 2016. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." *American Economic Journal: Economic Policy* 8 (3): 203–32.

Easterbrook, Frank H. 1983. "Criminal Procedure as a Market System." *The Journal of Legal Studies* 12 (2): 289–332.

Farrell, Jill. 2003. "Mandatory Minimum Firearm Penalties: A Source of Sentencing Disparity?" *Justice Research and Policy* 5 (1): 95-115.

Fischman, Joshua B. and Max M. Schanzenbach. "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, Travis W. and Tri Keah S. Henry. 2019. "Racial Disparities in Federal Sentencing Outcomes: Clarifying the Role of Criminal History." *Crime & Delinquency*.

Glaeser, Edward L., Daniel P. Kessler, and Anne Morrison Piehl. 2000. "What Do Prosecutors Maximize? An Analysis of Federalization of Drug Crimes." *American Law and Economics Review* 2 (2): 259-290.

Goncalves, Felipe and Steven Mello. 2018. "A Few Bad Apples? Racial Bias in Policing." Working paper.

Gordon, Sanford C. 2009. "Assessing Partisan Bias in Federal Public Corruption Prosecutions." *American Political Science Review* 103 (04): 534–54.

Grogger, Jeffrey and Greg Ridgeway. 2006. "Testing for Racial Profiling in Traffic Stops From Behind a Veil of Darkness." *Journal of the American Statistical Association* 101 (475): 878– 887.

Grossman, Gene M. and Michael L. Katz. 1983. "Plea Bargaining and Social Welfare." *American Economic Review* 73 (4): 749-757.

Haney, Craig. 2001. "The Psychological Impact of Incarceration: Implications for Post-Prison Adjustment." National Policy Conference.

Honold, Dan. 2014. "Note: Quantity, Role, and Culpability in the Federal Sentencing Guidelines." *Harvard Journal on Legislation* 51: 389-314.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933–959.

King, Ryan D. and Michael T. Light. 2019. "Have Racial and Ethnic Disparities in Sentencing Declined?" *Crime and Justice*.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Soren Pedersen, and Emmanuel Saez. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." 2011. *Econometrica* 79 (3): 651–692.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics* 8: 435–464.

Knorre, Alexey. 2017. "How to detect fraud in drug crime records? An analysis of 585,000 crime records in Russia." Working paper.

Knowles, John, Nicola Persico, and Petra Todd. 2001. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109 (1): 203–229.

Knox, Dean, Will Lowe, and Jonathan Mummolo. 2019. "The Bias Is Built In: How Administrative Records Mask Racially Biased Policing." Working paper.

Landes, William M. 1971. "An Economic Analysis of the Courts." *Journal of Law and Economics* 14 (1): 61-107.

Lee, David S. and Justin McCrary. 2017. "The Deterrence Effect of Prison: Dynamic Theory and Evidence." *Advances in Econometrics* 38.

Lynch, Mona. *Hard Bargains: The Coercive Power of Drug Laws in Federal Court.* New York: Russell Sage Foundation, 2016.

Mueller-Smith, Michael. 2015. "The Criminal and Labor Market Impacts of Incarceration." University of Michigan Working Paper.

Mustard, David B. 2001. "Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts." *Journal of Law and Economics* 44: 285–314.

Nyhan, Brendan and M. Marit Rehavi. 2017. "Tipping the Scales? Testing for Political Influence on Public Corruption Prosecutions." Working paper.

Oswald, Frederick, Gregory Mitchell, Hart Blanton, and Philip Tetlock. 2013. "Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies." *Journal of Personality and Social Psychology* 105 (2): 171-192.

Pfaff, John. *Locked In: The True Causes of Mass Incarceration and How to Achieve Real Reform*. New York: Basic Books, 2017.

Rehavi, M. Marit, and Sonja B. Starr. 2012. "Racial Disparity in Federal Criminal Sentences and Its Sentencing Consequences." Univ. of Michigan Law School Program in Law and Economics Working Paper Series. Working paper no. 12-002.

Rehavi, M. Marit, and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6):1320–54.

Reinganum, Jennifer F. 1988. "Plea Bargaining and Prosecutorial Discretion." *American Economic Review* 78 (4): 713-728.

Rhodes, William, Ryan Kling, Jeremy Luallen, and Christina Dyous. 2015. "Federal Sentencing Disparity: 2005-2012." Bureau of Justice Statistics Working Paper Series.

Robert Wood Johnson Foundation (RWJF). 2018. "Discrimination in America: Final Summary."

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3): 180–212.

Sallee, James M., and Joel Slemrod. 2012. "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." *Journal of Public Economics* 96 (11): 981-999.

Shermer, Lauren O'Neill and Brian D. Johnson. 2009. "Criminal Prosecutions: Examining Prosecutorial Discretion and Charge Reductions in U.S. Federal District Courts." *Justice Quarterly* 1-37.

Shively, Michael. 2005. "Study of Literature and Legislation on Hate Crime in America." National Institute of Justice.

Simon, Jonathan. 2007. *Governing Through Crime: How the War on Crime and American Democracy Created a Culture of Fear*. Oxford: Oxford Univ. Press

Sklansky, David Alan. 2018. "The Problems with Prosecutors." *Annual Review of Criminology* 1: 451–469.

Sloan, CarlyWill. 2019. "Racial Bias by Prosecutors: Evidence from Random Assignment." Working paper.

Spohn, Cassia, John Gruhl, and Susan 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, Seth. 2014. "The Cost of Racial Animus on a Black Candidate: Evidence using Google Search Data." *Journal of Public Economics* 118: 26-40.

Stuntz, William J. 2004. "Plea Bargaining and Criminal Law's Disappearing Shadow." *Harvard Law Review* 117 (8): 2548.

The Hamilton Project. 2016. "Twelve Facts about Incarceration and Prisoner Reentry." Economic Facts, The Hamilton Project, The Brookings Institution, Washington, DC

Traxler, Christian, Franz G. Westermaier, and Ansgar Wohlschlegel. 2018. "Bunching on the Autobahn? Speeding Responses to a 'Notched' Penalty Scheme." *Journal of Public Economics* 157: 78–94.

Ulmer, Jeffrey T., Megan C. Kurlychek, and John H. Kramer. 2007. "Prosecutorial Discretion and the Imposition of Mandatory Minimum Sentences." *Journal of Research in Crime and Delinquency* 44 (4): 427-458.

Ulmer, Jeffrey, Noah Painter-Davis, and Leigh Tinik. 2014. "Disproportional Imprisonment of Black and Hispanic Males: Sentencing Discretion, Processing Outcomes, and Policy Structures." *Justice Quarterly*.

US DOJ. 2012. "NIBRS Participation by State."

US DOJ. 2011. "The Department of Justice's International Prisoner Transfer Program."

USSC, 2015a. "Report to the Congress: Impact of the Fair Sentencing Act of 2010."

USSC. 2015b. "United States Sentencing Commission (USSC) Guidelines Manual."

USSC. 2017. "Demographic Differences in Sentencing: An Update to the 2012 Booker Report."

Webb, Dan K. and Scott F. Turow. 1982. "The Prosecutor's Function in Sentencing." *Loyola U. Chicago Law Journal* 13 (4): 641-668.

West, Jeremy. 2018. "Racial Bias in Police Investigations." Working paper.

Wright, Ronald. 2006. "Federal or State? Sorting as a Sentencing Choice." *Criminal Justice* 21 (2): 16-21.

Xu, Frank, Brian Nosek, Anthony Greenwald, Kate Ratliff, Yoav Bar-Anan, Emily Umanksy, Mahzarin Banaji, Nicole Lofaro, Colin Smith, Jordan Axt. 2019. "Project Implicit Demo Website Datasets."

Yang, Crystal S. 2016. "Resource Constraints and the Criminal Justice System: Evidence from Judicial Vacancies." *American Economic Journal: Economic Policy* 8 (4): 289–332.

Tables and Figures

Table 1. Summary Statistics for USSC Sentencing Data.

	1999-2010	2011-2015
Black or Hispanic	0.921	0.939
	(0.270)	(0.239)
Age (in years)	31.187	34.166
	(8.517)	(8.748)
Male	0.915	0.916
	(0.279)	(0.277)
College or more	0.126	0.148
	(0.332)	(0.355)
High school or more	0.509	0.598
	(0.500)	(0.490)
Not US citizen	0.046	0.033
	(0.209)	(0.178)
Weapon involved	0.262	0.296
	(0.440)	(0.456)
Number of other current offenses	1.606	1.720
	(1.427)	(1.735)
Criminal history points	5.713	6.512
	(5.474)	(5.586)
Drug weight (in grams)	102.530	116.968
	(156.957)	(169.892)
Sentence (in years)	9.294	7.807
	(7.057)	(5.833)
Observations	47,439	9,445

Notes. The table above describes defendants found in the USSC sentencing data pre- and post-2010. The mean value of each variable is reported with standard deviations in parentheses. The statistics above are derived from the cleaned USSC data in which the following cases are removed: cases with missing drug weight values (including those cases with weights coded as a range), cases with reported problems in the drug weight variables, cases where judges change or do not accept the findings of fact for drug weights, and cases at and above 1000g.

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

	Pr(280-290g Crack-Cocaine Recorded)		
	(1)	(2)	(3)
After 2010	0.0347***		0.0754***
	(0.00204)		(0.0132)
After 2010 x White		0.0125**	
		(0.0053)	
After 2010 x Black or Hispanic		0.0360***	
		(0.0021)	
Constant	0.0051***	0.0032***	0.00333***
	(0.0003)	(0.0010)	(0.00118)
P-value: W (White) = BH (Black or Hispanic)	-	0.0000	-
Trial Cases Only	No	No	Yes
Observations	56,884	52,745	2,823

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes on sample selection. The row "P-value: W (White) = BH (Black or Hispanic)" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." In the remaining tables, I abbreviate the label to "P-value: W= BH." Specifications with the race and after 2010 interactions also include a dummy variable equal to one for black and Hispanic offenders and equal to zero for white offenders. Coefficients are estimated from the following regression for Column 1:

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

and the following regression for Column 2:

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

Column 3 re-estimates equation (1) excluding cases that end in a plea deal (i.e. trial cases only). I do not re-estimate equation (2) on the trial-only sample because there are zero white offenders with 280-290g in trial cases after 2010.

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

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

Panel A. Analysis of Changes in the 0-100g Range.

	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	0.0172***	-0.0711***	0.0358***	-0.0061**	-0.0089**
	(0.0038)	(0.0048)	(0.0039)	(0.0028)	(0.0036)
Constant	0.1139***	0.2920***	0.1099***	0.0714***	0.1232***
	(0.0015)	(0.0021)	(0.0014)	(0.0012)	(0.0015)
Predicted Change from Conceptual Model	Increase	Decrease	Ambiguous	Decrease	Decrease
Observations	56,884	56,884	56,884	56,884	56,884

Panel B. Analysis of Changes in the 100-1000g Range.

	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
	(6)	(7)	(8)	(9)	(10)
After 2010 (Actual Change)	-0.0152***	0.0347***	0.0055**	0.0019	0.0062***
	(0.0043)	(0.0020)	(0.0024)	(0.0017)	(0.0020)
Constant	0.1929***	0.0051***	0.0439***	0.0214***	0.0263***
	(0.0018)	(0.0003)	(0.0009)	(0.0007)	(0.0007)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	56,884	56,884	56,884	56,884	56,884

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 sample selection. The predicted change from the conceptual model of prosecutor behavior in Section II.B is displayed in the row labeled "predicted change from conceptual model." Coefficients are estimated from the following regression for each range:

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

Tables A9f-g display versions of this table with race interactions. Tables A9a-e display versions of this table with time trend interactions.
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Racial Difference in Shifting from 50g Compared to Shifting to 280g

	Pr(50-60g)	Pr(280-290g)
	(1)	(2)
After 2010 x Black or Hispanic	-0.0066**	0.0360***
	(0.0029)	(0.0021)
After 2010 x White	-0.0006	0.0125***
	(0.0111)	(0.0053)
Constant	0.0653***	0.0032*
	(0.0042)	(0.0010)
P-value: BH = W	0.6000	0.0000
Observations	52,745	52,745

Notes. Robust standard errors 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 about sample selection. Coefficients are estimated from the following regression for each range:

$$(4) \quad (\text{Charged X-Yg})_{it} = \alpha_0 + \beta_1 \text{After2010}_{it} + \epsilon_{it}$$

Adding the coefficient in column (1) for black and Hispanic offenders to the coefficient in column (2) for black and Hispanic offenders yields a new coefficient of 0.0293. This coefficient is still larger than the coefficient in column (2) for white offenders and the two are statistically different at the one percent level (p-value = 0.0084).
*** p<0.01, ** p<0.05, * p<0.1

Table 5. Bunching Analysis for Potential Mechanisms

Panel A. Analysis of Bunching in State Convictions and in Drug Seizures

	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)	Pr(280-290g)	Pr(280-290g)
	(1)	(2)	(3)	(4)	(5)
After 2010	0.00005		-0.0002***		-0.0006***
	(0.0005)		(.0001)		(0.0002)
After 2010 x White		0.0004		-0.0001	
		(0.0011)		(0.0001)	
After 2010 x Black or Hispanic		0.0002		-0.0003***	
		(0.0005)		(0.0001)	
Constant	0.0051***	0.0085***	0.0004***	0.0002***	0.0010***
	(0.0003)	(0.0005)	(0.00005)	(0.0001)	(0.0001)
Data Analyzed	FL Convictions	FL Convictions	Drug Seizures, NIBRS	Drug Seizures, NIBRS	Drug Evidence, DEA STRIDE
Drugs Included	Cocaine, all types	Cocaine, all types	Crack-cocaine	Crack-cocaine	Cocaine, all types
P-value: W = BH	-	0.8148	-	0.2382	-
Observations	214,573	214,573	203,700	191,774	100,306

Panel B. Analysis of Bunching in Prosecutor Case Files and Final Sentencing

	Pr(280-290g)	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)	Pr(280-290g)
	(6)	(7)	(8)	(9)	(10)
After 2010	0.0783***	0.0408***		0.0347***	
	(0.00561)	(0.0126)		(0.00204)	
After 2010 x White			0.0031		0.0125**
			(0.0292)		(0.0053)
After 2010 x Black or Hispanic			0.0447***		0.0360***
			(0.0130)		(0.0021)
Constant	0.0039***	0.1096***	0.1242***	0.0051***	0.0032***
	(0.0004)	(0.0072)	(0.0156)	(0.0003)	(0.0010)
Data Analyzed	EOUSA Case Management System	USSC Sentencing, FL only	USSC Sentencing, FL only	USSC Sentencing	USSC Sentencing
Drugs Included	Crack-cocaine	Cocaine, all types	Cocaine, all types	Crack-cocaine	Crack-cocaine
P-value: W = BH	-	-	0.1566	-	0.0000
Observations	19,363	6,856	6,856	56,884	52,745

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

Table 6a. Offender Drug-Holding Behavior by Race, After Fair Sentencing Act in 2010

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

Notes. Robust standard errors estimated jointly by seemingly unrelated regression in parentheses. This analysis uses the weights of seized drugs reported to the FBI through the National Incident Based Reporting System. Ethnicity is not consistently recorded in NIBRS over this time period. As such, I refer to offenders as black or white, omitting the Hispanic label used in previous analyses. Columns 1-3 show the relationship between race of offender and drug weight seized, in general. Column 4 shows how the weight of an offender's seized drugs changes by race after 2010. Columns 5-11 show how the probability an offender's seized drugs are in a certain bin changes by race after 2010. All specifications include state fixed effects and controls for age and sex. The row "P-value: W= B" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black." Coefficients in column 1 are estimated from the following regression:

$$(5) \quad Weight_i = \alpha_0 + \beta_1 Black_i + X_i + Z_s + \epsilon_i$$

where $Weight_i$ is the weight of the drugs seized, $Black_i$ is an indicator of whether the offender is recorded as black or white, X_i includes offender age and sex, and Z_s is a vector of state fixed effects. The coefficients in column 2 are estimated from the same specification with a dummy variable for the 280-290g range as the dependent variable. Coefficients in column 3 are estimated from the following regression:

$$(6) \quad Weight_{it} = \alpha_0 + \beta_1(\text{Black} \times \text{After2010})_{it} + \beta_2(\text{White} \times \text{After2010})_{it} + X_i + Z_s + \epsilon_{it}$$

The coefficients in columns 4-10 are estimated from the same specification with dummy variables for the range of interest as the dependent variable.
*** p<0.01, ** p<0.05, * p<0.1

Table 6b. Drug Use and Drug Selling After the Fair Sentencing Act

	Ever Use Crack (1)	Sold Drugs in Past Year (2)	Use Crack & Sold Drugs (3)
After 2010 x White	0.0019**	-0.0009**	-0.0007***
	(0.0009)	(0.0005)	(0.0002)
After 2010 x Black or Hispanic	-0.0053***	-0.0031***	-0.0010***
	(0.0015)	(0.0009)	(0.0003)
Black or Hispanic	0.0033***	0.0039***	-0.0009***
	(0.0012)	(0.0007)	(0.0003)
Constant	0.0342***	0.0145***	0.0037***
	(0.0005)	(0.0007)	(0.0001)
Observations	763,335	762,322	762,054
P-value: W = BH	0.0000	0.0257	0.3350

Notes. Robust standard errors in parentheses. This analysis uses data from the National Survey on Drug Use and Health. Column 1 shows that the fraction of respondents answering "yes" to the question, "have you ever, even once, used crack-cocaine?" does not increase after 2010. Column 2 shows that the fraction of respondents answering a number greater than zero to the question, "how many times have you sold illegal drugs in the past 12 months?" does not increase after 2010. Column 3 shows that the fraction of people answering yes to both of these questions does not increase after 2010. All specifications use year-specific sampling weights. The row "P-value: W= BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." Coefficients in are estimated from the following regression:

$$\begin{aligned}(7)\quad \text{Outcome}_{it} &= \alpha_0 + \beta_1(\text{BlackOrHispanic}\times\text{After2010})_{it} \\ &+ \beta_2(\text{White}\times\text{After2010})_{it} + \text{BlackOrHispanic}_i + \epsilon_{it}\end{aligned}$$

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

Table 7. 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	-0.1794***	0.2170***	-0.0376
	(0.0629)	(0.0393)	(0.0461)
Constant	0.9184***	-	0.0816*
	(0.0435)	-	(0.0435)
Observations	989	989	989
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Atty. Bunches at 50-60g Pre-2010	-0.0785***	0.0575***	0.0211
	(0.0254)	(0.0172)	(0.0168)
Constant	0.9359***	0.0233**	0.0408***
	(0.0170)	(0.0105)	(0.0133)
Observations	1,135	1,135	1,135

Notes. Standard errors clustered at the prosecutor level and estimated jointly by seemingly unrelated regression in parentheses. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

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

where AttyBunchesAt280g is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). These regressions are restricted to post-2010 cases (for columns 1-3) and to prosecutors with 10+ cases post-2010. Note, column (2) is a mechanical relationship, hence the missing standard error. Table A22 shows that this result is robust to using leave-out-means to classify bunching attorneys. Coefficients in panel B are estimated from the following regression for each range:

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

where AttyBunchesAt50g is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 50g definition (i.e. the fraction of their cases that are from 50-60g is above the average fraction of 50-60g cases post-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 50-60g is at or below the average fraction of 50-60g cases post-2010). These regressions are restricted to post-2010 cases (for columns 4-6) and to prosecutors with 10+ cases pre-2010.

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

Table 8. Change in Bunching by Prosecutors after *Alleyne v. United States* Decision

	Pr(Case Recorded with 280-290g)			
	(1)	(2)	(3)	(4)
After June 17th, 2011-2016	0.0070	-0.0049	0.0041	-0.0206
	(0.0260)	(0.0284)	(0.0295)	(0.0406)
After June 17th, 2013	-0.1740**	-0.1518*	-0.1433	-0.1289
	(0.0813)	(0.0920)	(0.0935)	(0.1246)
Constant	0.1620	0.1626	0.1576	0.2093
	(0.1520)	(0.1519)	(0.1520)	(0.1776)
Bandwidth	±150 days	±130 days	±120 days	±60 days
Observations	1,937	1,672	1,513	754

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

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

where AfterJune17 is a dummy variable equal to one for cases received after June 17th in each year, DaysFrom, the running variable, is the date the case was received centered at zero on June 17th, and Year2013 is equal to one for cases received in 2013 (the year *Alleyne* is decided). In addition, all specifications above include day-of-week fixed effects, D_{it}, for the day the case is received. The ±130 day bandwidth is selected from the Imbens-Kalyanaraman optimal bandwidth procedure for the year 2013. Figure 4 shows graphical evidence of the discontinuity in bunching around June 17, 2013. Figure A21 shows further robustness checks.
*** p<0.01, ** p<0.05, * p<0.1

Table 9. Degree of Bunching Post-2010 by Race and Offender Characteristics.

	Pr(280-290g)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After ’10 x White (W)	0.0171**	0.0065	0.0143	0.0129**	0.0160**	0.0103*	0.0149*	-0.0018**	0.0085
	(0.0068)	(0.0063)	(0.0087)	(0.0062)	(0.0071)	(0.0058)	(0.0076)	(0.0009)	(0.0095)
After ’10 x Black or Hispanic (BH)	0.0363***	0.0235***	0.0424***	0.0303***	0.0452***	0.0306***	0.0471***	0.0088***	0.0156***
	(0.0023)	(0.0072)	(0.0037)	(0.0024)	(0.0036)	(0.0025)	(0.0173)	(0.0015)	(0.0040)
After ’10 x W x Char.	-0.0207***	0.0109	-0.0024	-0.0015	-0.0095	0.0089	-0.0074	0.0283***	0.0067
	(0.0072)	(0.0100)	(0.0109)	(0.0120)	(0.0107)	(0.0135)	(0.0110)	(0.0109)	(0.0123)
After ’10 x BH x Char.	-0.0042	0.0131*	-0.0102**	0.0191***	-0.0163***	0.0157***	-0.0188	0.0686***	0.0250**
	(0.0061)	(0.0076)	(0.0046)	(0.0051)	(0.0044)	(0.0047)	(0.0183)	(0.0052)	(0.0118)
Constant	0.0032***	0.0022	0.0031**	0.0033***	0.0013*	0.0036***	0.0031**	0.0018**	0.0052**
	(0.0011)	(0.0015)	(0.0014)	(0.0011)	(0.0008)	(0.0012)	(0.0014)	(0.0009)	(0.0020)
Characteristic	College	Male	Above Med. Age	Weapon	Above Med. Crim. Hist. Points	Above Med. # of Other Counts	State Above Med. % of Black and Hispanic Cases	Conspiracy Charge	State Above Med. Racial Animus
P-value: W = BH	0.0074	0.0764	0.0031	0.0085	0.0002	0.0012	0.0885	0.0000	0.5114
P-value: W+Char. = BH+Char.	0.0000	0.0177	0.0043	0.0007	0.0078	0.0352	0.0183	0.0000	0.0440
Observations	52,389	49,049	52,712	52,233	52,725	52,742	52,692	52,745	51,679

Notes. Robust standard errors in parentheses for columns 1-6. Standard errors clustered at the state level are in parentheses for columns 7 and 9. “Characteristic” or “Char.” represents a dummy variable that is an offender or case characteristic. 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 sample selection. The row “P-value: W = BH” reports the p-value from a test of the null hypothesis that the coefficient on “After 2010 x White” is equal to the coefficient on “After 2010 x Black or Hispanic.” The row “P-value: W+Char. = BH+Char.” reports the p-value from a test of the null hypothesis that the combined coefficients on “(After 2010 x White)+(After 2010 x White x Characteristic)” is equal to the combined coefficients on “(After 2010 x Black or Hispanic)+(After 2010 x Black or Hispanic x Characteristic).” Male is equal to one if the offender is male and equal to zero if not. Above median age is equal to one if the offender is above the median age for offenders and equal to zero if not. Weapon is equal to one if the offense involves a weapon and equal to zero if not. Above median crim. hist. points is equal to one if the offender has a criminal history score above the median criminal history score for offenders and equal to zero if not. Above the median # of other counts is equal to one if the offender has above the median number of other criminal counts for offenders and equal to zero if not. Column 7 examines differences in bunching for offenders convicted in states with above/below the median fraction of black and Hispanic cases. Column 8 tests for differences in bunching for offenders with a “drug conspiracy” charge versus those without. The final column examines differences in bunching for offenders convicted in states with above/below the median level of racial animus. The coefficients in columns 1-9 are estimated from the following regression:

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

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

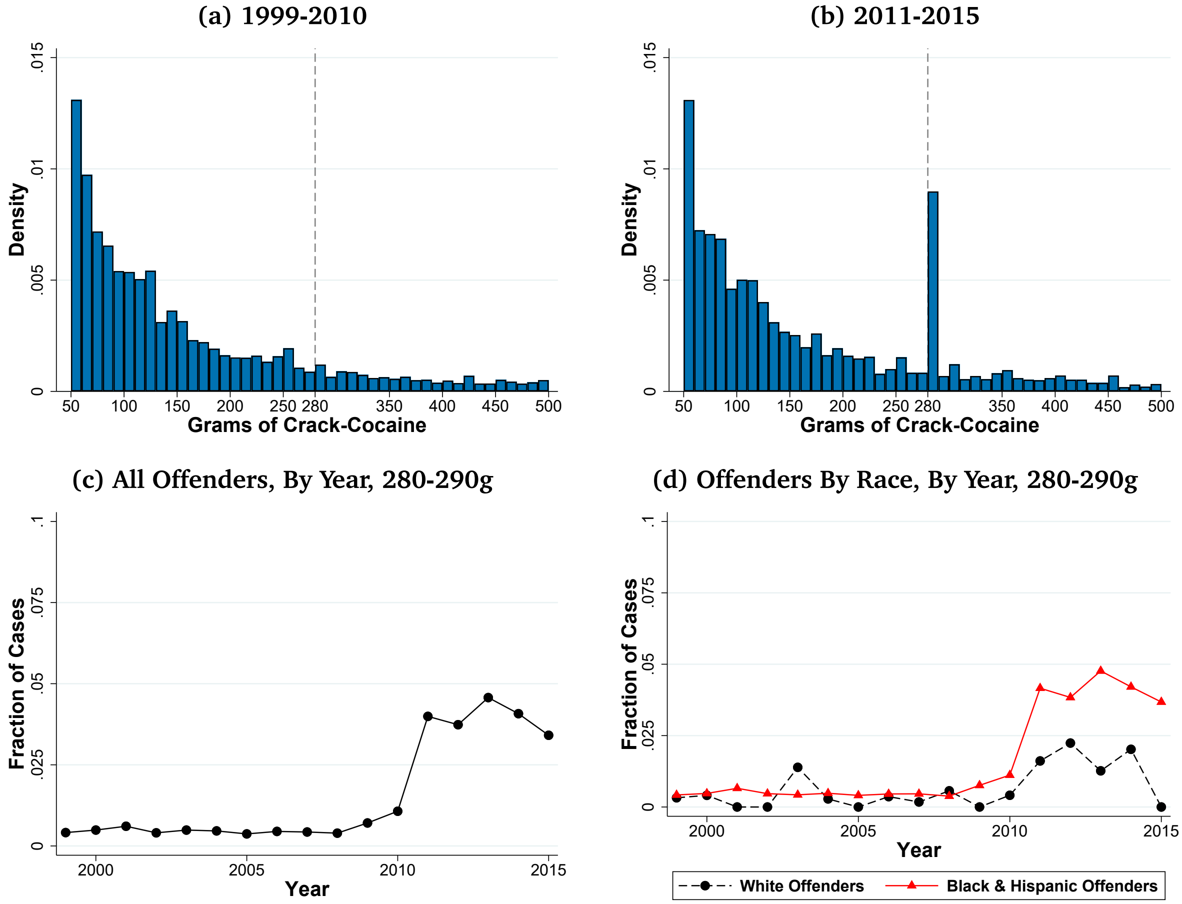
Table 10. **Robustness Tests for Relationship between Racial Animus and the Racial Disparity Bunching at 280g**

	Pr(280-290g)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After '10 x W x Above Med. Animus	0.0067	-0.0033	0.0063	-0.0047			
	(0.0123)	(0.0228)	(0.0128)	(0.0245)			
After '10 x BH x Above Med. Animus	0.0250**	0.0267**	0.0269**	0.0279**			
	(0.0118)	(0.0108)	(0.0124)	(0.0108)			
After '10 x W x Continuous Animus					0.0001	0.0008	
					(0.0004)	(0.0008)	
After '10 x BH x Continuous Animus					0.0007	0.0015***	
					(0.0005)	(0.0004)	
After '10 x IAT-Lawyers							-0.0075
							(0.0095)
After '10 x BH x IAT-Lawyers							0.0155
							(0.0105)
Constant	0.0052**	-0.0334	0.0040*	-0.0282	0.0099**	0.0111	0.0037
	(0.0020)	(0.0295)	(0.0023)	(0.0300)	(0.0044)	(0.0081)	(0.0059)
Other Controls Included	None	Offender Controls	District Economic Controls	Offender + District Controls	None	None	State x After 2010 x Race FEs
Sample Restrictions	None	None	None	None	None	Outliers Removed	None
Observations	51,679	51,679	47,692	47,692	51,679	49,188	51,679

Notes. Standard errors clustered at the state level in parentheses for columns 1-6. Standard errors clustered at the district level are in parentheses for column 7. See Table 1 for notes on sample selection. The first four columns examine differences in bunching for offenders convicted in states with above/below the median level of racial animus. Column 1 reports this result with no additional controls; column 2 introduces individual controls (college, male, age, criminal history, and state caseload) interacted with the after 2010 by race dummy variables; column 3 introduces district controls for economic characteristics (median household income in 2016, non-white share of population in 2010, population density in 2010, fraction with college in 2010, poor share in 2010, log of wage growth for high school graduates, black-white and Hispanic-white differences in incarceration and income conditional on parent income rank at the 25th percentile, job density in 2013, and annual job growth from 2004-2013) interacted with the after 2010 by race dummy variables; column 4 combines all controls from columns 2-3. Column 5 examines the relationship between animus and bunching using the continuous measure of animus from Google Trends, the p-value is less than 0.2 and the coefficient is several times larger than the coefficient for white offenders. Column 6 re-runs column 5 with outlier states (states with animus above the 99th percentile or below the 1st percentile) removed. Column 7 introduces a district level measure of animus, the implicit association test scores for lawyers (and other legal-service workers) aggregated to the district level. Since the measure is at the district level, I include state fixed effects interacted with the after 2010 by race dummy variables. The estimate is identified from within-state variation in the IAT-animus measure, and the p-value on the estimate is 0.14. The IAT measure is scaled to the median difference between the minimum and maximum score in states, meaning a one unit increase is approximately equivalent to moving from the minimum score in a state to the maximum score.

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

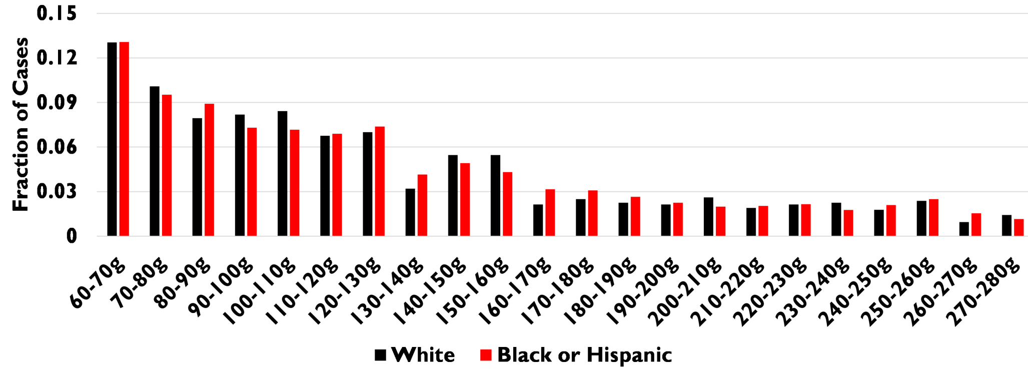
Figure 1. Changing Distribution of Drug Amounts Around 280g Pre- and Post-2010.



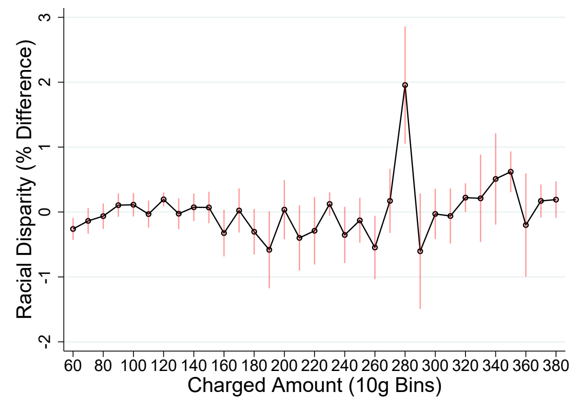
Notes. Panels (a) and (b) plot the distribution of drug amounts recorded in federal crack-cocaine sentences starting at 50 grams and ending at 500 grams for 1999-2010 (when the mandatory minimum threshold was 50g) and 2011-2015 (when it was 280g). Panels (c) and (d) display the fraction of crack-cocaine cases with 280-290g by year, in general and by race. The denominator in panel (c) is all crack-cocaine cases under 1000g. The denominators in panel (d) are all crack-cocaine cases under 1000g, by race. Histograms showing the full density from 0-500g are in Figures A3a-b. Figures 1c-d with confidence intervals are in Figures A3c-d.

Figure 2. Testing for Conditional Racial Disparity in Bunching

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



(b) Shifting from 60-380g by Race

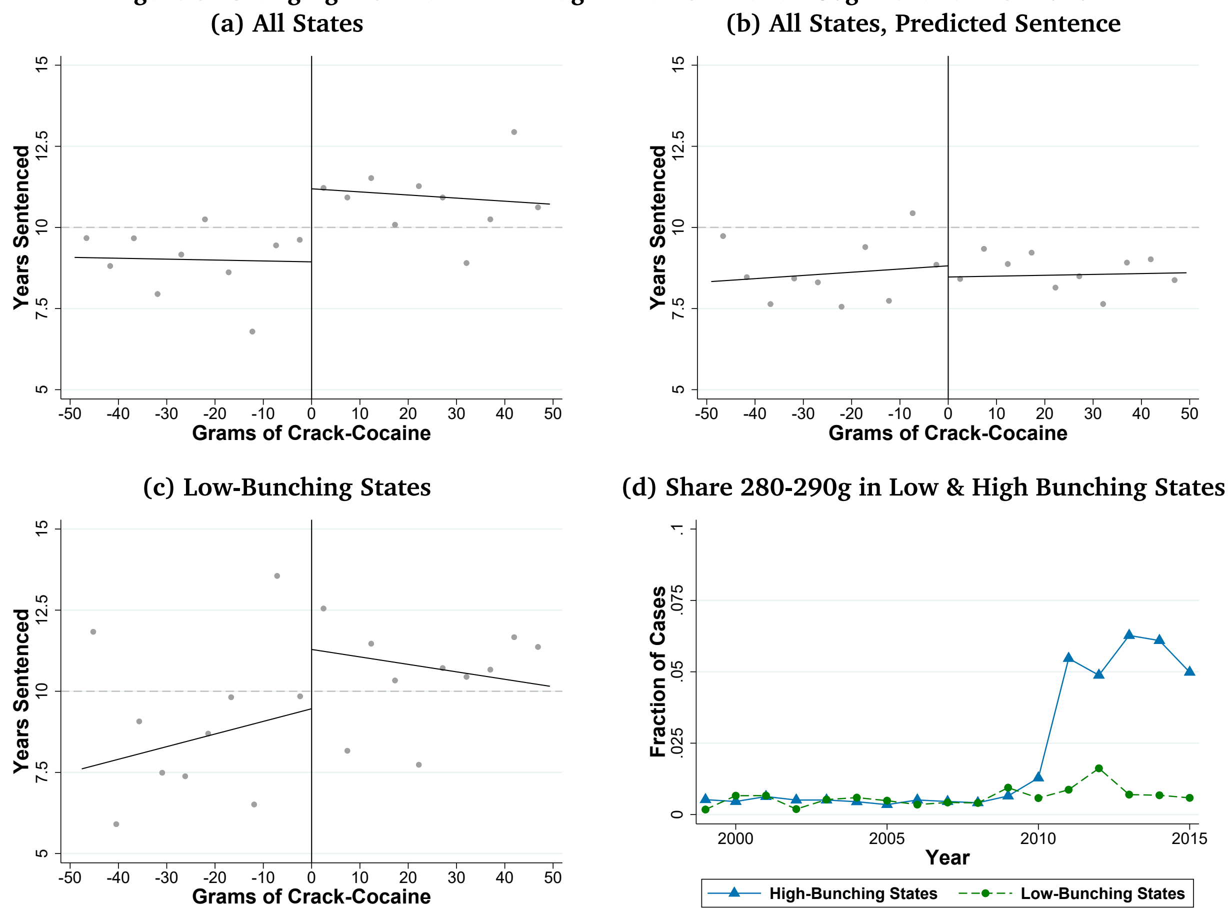


Notes. Panel (a) plots the distribution of charged amounts pre-2010 from 60-280g. A Kolmogorov-Smirnov test of the equality of the distributions by race fails to reject the null that the distributions are equal (p-value=0.788). Panel (b) plots the coefficient δ^X for each 10g bin starting at X divided by the share of cases in each 10g bin.

$$\text{(12)} \quad (\text{Charged X-Yg})_{it} = \alpha + \delta^X(\text{After2010}\times\text{BlackOrHispanic})_{it} + \gamma\text{After2010}_{it} + \lambda\text{BlackOrHispanic}_i + \epsilon_{it}$$

The plot shows these estimates for amounts from 0-380g, at higher amounts the estimates are more noisy. Figure A8 shows the estimates up to 1000g.

Figure 3. Changing Distribution of Drug Amounts Around 280g Pre- and Post-2010.

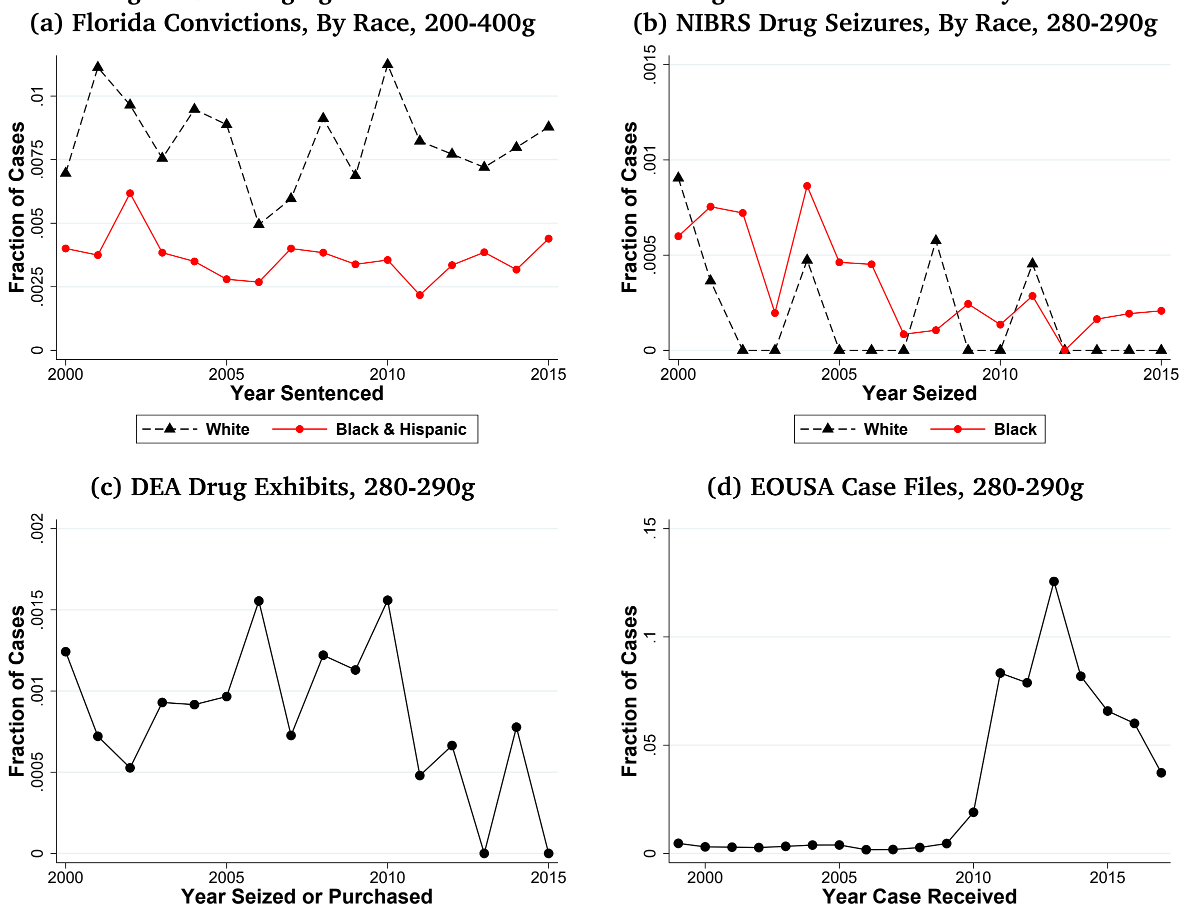


Notes. Figure 3a plots the average sentence (within each 5g bin) from 230-330g for cases sentenced after 2010. A linear fit is estimated on each side of the 280g threshold. The estimated sentencing discontinuity is about 2.25 years ($se = 0.85$). Figure 3b is the same plot but using predicted sentence from a model of sentencing and offender characteristics using pre-2010 data. There is no discontinuity in this figure, suggesting that offenders bunched at 280g are not negatively selected on characteristics that would increase sentence length in the absence of the threshold. Figure 3c is the same plot but limited to the subset of states that have low-levels of bunching. The estimated discontinuity is about 2.00 years ($se = 1.73$). Figure 3d plots the share of cases with 280-290g by year for low- and high-bunching states. The coefficients described above are estimated from the regression:

$$\text{(13)} \quad \text{Sentence}_i = \alpha_0 + \beta_1 \text{Amount}t_i + \delta_1 \text{Above280}_i + \phi_1(\text{Amount} \times \text{Above280}_i) + \epsilon_i$$

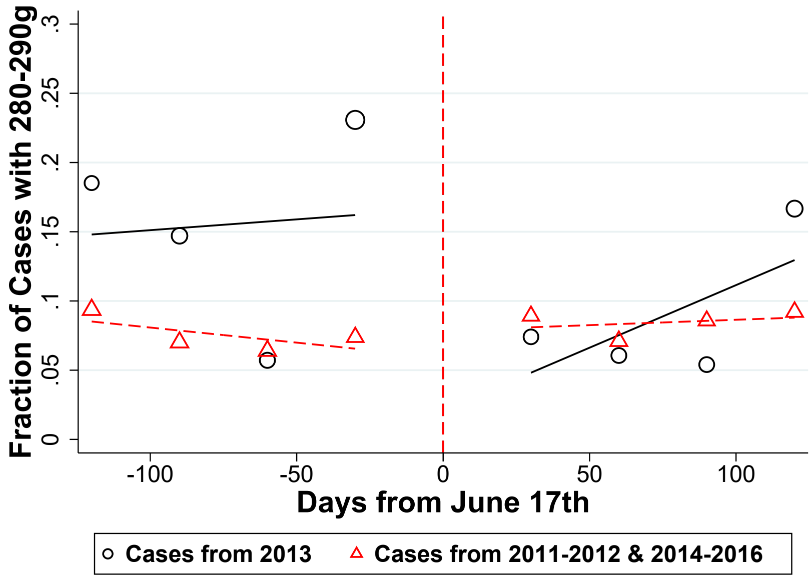
δ_1 is the estimated discontinuity (reported in the preceding notes) in sentencing due to crossing the mandatory minimum threshold.

Figure 4. Changing Fraction of Cases at Various Stages of Criminal Justice System



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

Figure 5. Additional Evidence of Prosecutorial Discretion in Bunching, Alleyne Results and Movers Results



Notes. Panel (a) plots the fraction of cases with 280-290g in each 30-day bin for 120 days before and 120 days after June 17th. The black circles show the fraction of cases in each bin for 2013 and the red triangles show the average fraction of cases in each bin for 2011-2012 and 2014-2016. The solid black line shows a linear fit on each side of the June 17, 2013 and the dashed red line shows a linear fit on each side of June 17 for all other years. The scatter plot symbols are weighted by the total number of cases in each bin. The estimated discontinuity is $\delta = -0.1433$ and $se = 0.0935$ and is estimated from the following regression:

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

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 is equal to zero if case **i** is received before June 17th of year **t** but after January 1st of year **t**. $DaysFrom_{it}$ is the number of days from June 17th that case **i** is received, and $Year2013_{it}$ is equal to one if case **i** is received in 2013 and is equal to zero if it is received in 2011-2012 or 2014-2016. D_{it} represents day-of-week fixed effects. The coefficient β_1 is the average discontinuity in the fraction of cases with 280-290g after June 17 from 2011-2016. The coefficient δ_1 is the discontinuity that is specific to June 17, 2013–the date of the *Alleyne* decision.

Appendix A. Additional Analyses

Table A1. Summary Statistics for FL, NIBRS, and DEA Records

	Pre-2010	Post-2010	Observations
Panel A. Cocaine Felony Convictions in FL			
200-400g	0.00474	0.00432	214,573
	(0.0687)	(0.0656)	
28-200g	0.0405	0.0473	214,573
	0.197	(0.212)	
Missing drug weight	0.945	0.936	214,573
	(0.228)	(0.245)	
Black or Hispanic	0.771	0.789	214,573
	(0.420)	(0.408)	
Panel B. NIBRS Drug Seizures, Balanced Panel			
Weight (g)	10.33	7.76	203,700
	(46.19)	(44.87)	
280-290g	0.000360	0.000141	203,700
	(0.0190)	(0.0119)	
Black	0.737	0.746	191,774
	(0.440)	(0.435)	
Male	0.837	0.834	192,721
	(0.370)	(0.372)	
Panel C. DEA Drug Seizures			
Weight (g)	78.28	67.28	100,306
	(188.83)	(176.54)	
280-290g	0.00102	0.000428	100,306
	(0.0319)	(0.0207)	
Seized (vs. Purchased)	0.529	0.544	100,302
	(0.499)	(0.498)	
Price per gram (median)	42.02	47.62	37,820

Notes. The table above describes offenders found in the FL inmate database, the NIBRS drug seizure records, and the DEA drug exhibit data pre- and post-2010 (the DEA data actually describes the drugs themselves, not the offenders). The mean value of each variable is reported with standard deviations in parentheses. Observation counts are displayed separately for each variable. The statistics above are derived from the cleaned data in which the following cases are removed for NIBRS and DEA: cases with drug weights above 1000g. Weight is the weight of the drugs in grams recorded. 280-290g is a dummy variable equal to one when the weight is from 280-290g and zero when it is from 0-280g and 290-1000g, and missing when it is missing. The 200-400g and 28-200g variables follow the same logic. Missing drug weight is equal to one when the drug weight is missing. "Seized (vs. Purchased)" is equal to one if the DEA obtained the drug exhibit from a seizure versus an undercover purchase. The median price per gram is reported after removing outliers above the 95th percentile and below the 5th percentile.

Table A2. Summary Statistics for EOUSA Prosecutor Case Files

	Pre-2010	Post-2010	Observations
Weight (g)	72.500	97.966	19,363
	(135.219)	(162.538)	
280-290g	0.004	0.082	19,363
	(0.062)	(0.274)	
280-290g, Missing = 0	0.002	0.026	49,342
	(0.040)	(0.158)	
50-60g	0.210	0.082	19,363
	(0.408)	(0.274)	
50-60g, Missing = 0	0.086	0.026	49,342
	(0.280)	(0.158)	
Missing drug weight	0.593	0.686	49,342
	(0.491)	(0.464)	
Only Federal Law Enforcement Involved	0.642	0.647	48,501
	(0.479)	(0.478)	
Any Federal Law Enforcement Involved	0.737	0.713	48,501
	(0.440)	(0.452)	
Lead Charge = Conspiracy	0.212	0.217	46,335
	(0.409)	(0.412)	

Notes. The table above describes defendants found in the EOUSA prosecutor case management data pre- and post-2010. The mean value of each variable is reported with standard deviations in parentheses. Observation counts are displayed separately for each variable since some fields in this data are missing much more often than others. The statistics above are derived from the cleaned data in which the following cases are removed: cases with drug weights above 1000g. Weight is the weight of the drugs in grams recorded in the case management system. 280-290g is a dummy variable equal to one when the weight is from 280-290g, zero when it is from 0-280g and 290-1000g, and missing when it is missing.. "280-290g, Missing=0" is a dummy variable equal to "280-290g" but coded equal to zero when the weight field is missing. The 50-60g variables follow the same logic. Missing drug weight is equal to one when the drug weight is missing. "Only Federal Law Enforcement" is equal to one when the agency recorded as sending the case is strictly federal (i.e. DEA, FBI, or ATF) and equal to zero otherwise. "Any Federal" is equal to one if the agency sending the case has any federal involvement (i.e. "Joint DEA and state/local task force") and equal to zero otherwise. "Lead Charge = Conspiracy" is equal to one when the lead charge for the case is a drug conspiracy charge.

Table A3. Result Robust to Other Drug Weight Sample Restrictions

	Pr(280-290g Crack-Cocaine)				
	(1)	(2)	(3)	(4)	(5)
After 2010 x White	0.0119**	0.0115**	0.0115**	0.0844***	0.0258**
	(0.0050)	(0.0049)	(0.0049)	(0.0131)	(0.0116)
After 2010 x Black or Hispanic	0.0345***	0.0329***	0.0328***	0.1186***	0.0718***
	(0.0021)	(0.0020)	(0.0020)	(0.0040)	(0.0042)
Constant	0.0031***	0.0030***	0.0030***	0.0034***	0.0088***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0027)
P-value: W = BH	0.0000	0.0000	0.0001	0.0127	0.0002
Sample Restriction	0-2500g	0-25000g	No Restriction	0-1000g	50-1000g
Includes Weights Coded as a Range	No	No	No	Yes	No
Observations	55,729	58,116	58,645	59,677	24,905

Notes. Robust standard errors in parentheses. 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." Columns 1-3 include outliers to varying extents. Column 4 reports results when the sample includes quantities coded as a range (in this analysis, the lower bound of the range is used). Column 5 excludes drug weights below 50g (i.e. excluding weights close to the 5-year mandatory minimum pre- and post-2010).

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

Table A4. Result Robust to Various Sample Restrictions

	Pr(280-290g Crack-Cocaine)					
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0314***		0.0336***		0.0304***	
	(0.0021)		(0.0021)		(0.0022)	
After 2010 x White		0.0125**		0.0128**		0.0128**
		(0.0053)		(0.0054)		(0.0054)
After 2010 x Black or Hispanic		0.0327***		0.0348***		0.0317***
		(0.0022)		(0.0022)		(0.0023)
Constant	0.0053***	0.0032***	0.0062***	0.0030**	0.0063***	0.0030**
	(0.0004)	(0.0010)	(0.0006)	(0.0015)	(0.0006)	(0.0015)
P-value: W = BH	-	0.0004	-	0.0002	-	0.0013
Hispanic Offenders Excluded	Yes	Yes	No	No	Yes	Yes
Post-2006 Data Only	No	No	Yes	Yes	Yes	Yes
Observations	47,763	47,763	25,893	25,846	23,241	23,241

Notes. Robust standard errors in parentheses. 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 “Post-2006 Data Only” is equal to “Yes” when the data is limited to cases brought to court from 2007-2015 (after the *Booker v. United States* Supreme Court case that made sentencing guidelines optional, excluding mandatory minimum guidelines). The row “Hispanic Offenders Excluded” is equal to “Yes” when Hispanic offenders are removed from the sample.

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

Table A5. Result Robust to Other Categorizations of Bunching

	Pr(280-300g)	Pr(280-320g)	Pr(280-380g)
	(1)	(2)	(3)
After 2010 x White	0.0154**	0.0146**	0.0137*
	(0.0061)	(0.0067)	(0.0083)
After 2010 x Black or Hispanic	0.0360***	0.0367***	0.0394***
	(0.0022)	(0.0025)	(0.0029)
Constant	0.0055***	0.0099***	0.0230***
	(0.0013)	(0.0017)	(0.0026)
P-value: W = BH	0.0016	0.0019	0.0033
Observations	52,745	52,745	52,745

Notes. Robust standard errors in parentheses. 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." Each column corresponds to a different definition of what it means for a case to be "bunched" above the mandatory minimum threshold. For the main results, I define a result as "bunched" if it is in the narrow range of 280-290g. In columns 1-3, I use alternative ranges: 280-300g, 280-320g, and 280-380g.

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

Table A6. Result Robust to Controls and Alternative Std. Errors.

	Pr(280-290g Crack-Cocaine)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After 2010	0.0347***		0.0348***		0.0345***		0.0327***		0.0322***	
	(0.0082)		(0.0081)		(0.0079)		(0.0068)		(0.0066)	
After 2010 x White		0.0125**		0.0130**		0.0136**		0.0118*		0.0138**
		(0.0058)		(0.0059)		(0.0062)		(0.0060)		(0.0066)
After 2010 x Black or Hispanic		0.0360***		0.0363***		0.0358***		0.0340***		0.0333***
		(0.0086)		(0.0086)		(0.0084)		(0.0073)		(0.0071)
Constant	0.0051***	0.0032***	0.0085***	0.0064**	0.0088**	0.0085**	0.0078*	0.0074*	0.0082**	0.0075**
	(0.0005)	(0.0010)	(0.0029)	(0.0031)	(0.0035)	(0.0037)	(0.0043)	(0.0044)	(0.0031)	(0.0033)
P-value: W = BH	-	0.0181	-	0.0184	-	0.0282	-	0.0286	-	0.0695
Offender Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Year Trend	No	No	No	No	No	No	Yes	Yes	Yes	Yes
State-specific Trends	No	No	No	No	No	No	No	No	Yes	Yes
Observations	56,826	52,692	51,813	51,746	51,813	51,746	51,804	51,737	51,804	51,737

Notes. Standard errors clustered at the state-level in parentheses. 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 "Offender Controls" indicates if the following offender-level controls are included: criminal history points, age, citizenship, number of current offense counts, whether a weapon was involved, and education. The rows "State Fixed Effects" and "Year Trend" indicate if the specification includes state fixed effects or a year trend as controls. The row "State-specific Trends" indicates if the specification includes state-specific linear trends. In all cases, there is a sharp increase in the fraction of cases with 280-290g after 2010 and a racial disparity in that increase by race.
*** p<0.01, ** p<0.05, * p<0.1

Table A7. Result Robust to Probit, Logit, and Poisson Models.

	Probit			Logit			Poisson			OLS		
	280-290g		280-380g	280-290g		280-380g	280-290g		280-380g	280-290g		280-380g
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
After 2010 x W	0.5747***	0.5606***	0.2046*	1.6031***	1.4119***	0.4804*	1.1208***	1.1615***	0.1102	0.0125**	0.0252**	0.0137*
	(0.1651)	(0.1840)	(0.1085)	(0.4518)	(0.4546)	(0.2498)	(0.4042)	(0.2758)	(0.5745)	(0.0053)	(0.0113)	(0.0083)
After 2010 x BH	0.8159***	0.9008***	0.3851***	2.0784***	2.0895***	0.8400***	2.1129***	2.1042***	0.8604	0.0360***	0.0710***	0.0394***
	(0.0337)	(0.0374)	(0.0235)	(0.0869)	(0.0878)	(0.0500)	(0.3645)	(0.2726)	(0.6351)	(0.0021)	(0.0042)	(0.0029)
Constant	-2.7258***	-2.3912***	-1.9948***	-5.7392***	-4.7715***	-3.7476***	3.5423***	2.6237***	3.6109***	0.0032***	0.0084***	0.0230***
	(0.0994)	(0.1102)	(0.0470)	(0.3020)	(0.3028)	(0.1138)	(0.3624)	(0.2202)	(0.3624)	(0.0010)	(0.0025)	(0.0026)
P-value: W = BH	0.1524	0.0701	0.1041	0.3015	0.1433	0.1580	0.0157	0.0007	0.3286	0.0000	0.0001	0.0033
Sample	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g
Observations	52,745	25,647	52,745	52,745	25,647	52,745	400	380	400	52,745	25,647	52,745

Notes. Robust standard errors in parentheses. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x W" is equal to the coefficient on "After 2010 x BH," where "W" is the "White" dummy variable and "BH" is the "Black or Hispanic" dummy variables (abbreviated for table space). In general, columns 1-3 estimate probit models, columns 4-6 estimate logit models, columns 7-9 estimate Poisson models (on binned data), and columns 10-12 estimate OLS (or linear probability) models. Columns 1, 4, 7, and 10 estimate the change in bunching at 280-290g after 2010 for all cases from 0-1000g. Columns 2, 5, 8, and 11 limit the sample to cases from 50-1000g (following column 5 of Table A3). Columns 3, 6, 9, and 12 extend the "bunching" definition to 280-380g (following column 3 of Table A5).

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

Table A8. Result Robust to Concerns about Selection Into/Out of Missing and Selection Into/Out of Other Drugs

	Pr(280-290g)			
	(1)	(2)	(3)	(4)
After 2010 x White	0.0583***	0.0242***	0.0005	0.0727***
	(0.0087)	(0.0059)	(0.0003)	(0.0032)
After 2010 x Black or Hispanic	0.0833***	0.0441***	0.0093***	0.2030***
	(0.0027)	(0.0021)	(0.0006)	(0.0031)
Constant	0.0033***	0.0024***	0.0004***	0.8680***
	(0.0009)	(0.0007)	(0.0001)	(0.0021)
P-value: W = BH	0.0063	0.0016	0.0000	0.0000
Drugs included	Crack-cocaine	Crack-cocaine	All	All
Dependent variable recoded to	Lower value of weight range	Upper value of weight range	Non-crack cases = 0	Non-crack cases = 1
Selection issue addressed	Into/out of missing weight	Into/out of missing weight	Into/out of other drugs	Into/out of other drugs
Observations	67,040	65,003	149,428	149,428

Notes. Robust standard errors in parentheses. 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 "Drugs included" indicates the type of drugs included in the analysis. In columns 1 and 2, I focus on the crack-cocaine sample to analyze how including missing exact weights (i.e. weights recorded as ranges) affects the results. In columns 3 and 4, I focus on the sample of all drugs to analyze how movement of cases into or out of other drug types affects the results. The row "Dependent variable recoded to" indicates how the dependent variable is recoded in each analysis. In column 1, the dependent variable is recoded as 1 if the lower bound of the weight range is between 280-290g and recoded as 0 otherwise. In column 2, it is recoded as 1 if the upper bound of the range is between 280-290g and recoded as 0 otherwise. Results are also robust to recoding all missings as (In 280-290)=0 or recoding all missings as (In 280-290)=1. In column 3, the dependent variable is recoded as 0 if the case is not a crack-cocaine case, and in column 4, it is recoded as 1 if the case is not a crack-cocaine case. Finally, the row "Selection issue addressed" indicates the type of selection issue being investigated in each column. In all columns, I find that the probability of being in the 280-290g range for crack-cocaine increases after 2010 and increases disproportionately for black and Hispanic offenders, regardless of selection into missing exact weights or other drug types.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A9. Difference-in-Difference Bunching Identification

	Pr(280-290g)				Pr(50-60g)	
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0011*	-0.0002				
	(0.0006)	(0.0011)				
After 2010 x Crack-cocaine	0.0336***	0.0127**				
	(0.0021)	(0.0054)				
After 2010 x Crack-cocaine x Black or Hispanic		0.0217***				
		(0.0059)				
Crack-cocaine	-0.0020***	-0.0042***	-0.0036**	0.0088	0.0151***	0.0210*
	(0.0005)	(0.0011)	(0.0016)	(0.0058)	(0.0053)	(0.0122)
Crack-cocaine x Black or Hispanic			0.0020	0.0229***	0.0108*	-0.0021
			(0.0017)	(0.0063)	(0.0057)	(0.0127)
Constant	0.0072***	0.0074***	0.0068***	0.0070***	0.0502***	0.0438***
	(0.0003)	(0.0006)	(0.0012)	(0.0026)	(0.0033)	(0.0065)
Drugs Included	All	All	Crack & Powder	Crack & Powder	Crack & Powder	Crack & Powder
Years Included	1999-2015	1999-2015	1999-2010	2011-2015	1999-2010	2011-2015
Observations	149,428	149,428	65,475	17,307	65,475	17,307

Notes. Robust standard errors in parentheses. Columns 1-2 compare crack-cocaine cases to all other drug cases. Specifically, they estimate the change in the probability a case is recorded with 280-290g after 2010 both for crack-cocaine and for other drugs. Column 1 does this in general and column 2 does this by race. This amounts to a difference-in-difference (pre- vs. post-2010 and crack vs. non-crack) estimation of the bunching (as opposed to the pre- vs. post-2010 difference that is the focus of the paper). Columns 3-6 apply this same design to estimate the probability of being recorded with 280-290g and 50-60g before and after 2010. These columns compare crack to powder cocaine alone since powder cocaine is a drug that never has a 50g mandatory minimum threshold.

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

Table A10a. Missing Mass in the Distribution of Drug Amounts by Race

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 x White	-0.0030	-0.1162***	0.0326**	-0.0006	0.0189
	(0.0179)	(0.0188)	(0.0149)	(0.0111)	(0.0143)
After 2010 x Black or Hispanic	0.0222***	-0.0696***	0.0341***	-0.0066**	-0.0100***
	(0.0039)	(0.0050)	(0.0041)	(0.0029)	(0.0038)
Constant	0.1971***	0.3242***	0.0968***	0.0653***	0.0965***
	(0.0068)	(0.0080)	(0.0050)	(0.0042)	(0.0050)
P-value: W = BH	0.1669	0.0164	0.9216	0.6000	0.0503
Observations	52,745	52,745	52,745	52,745	52,745

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)
	(1)	(2)	(3)	(4)	(5)
After 2010 x White	0.0028	0.0125**	0.0137	0.0099	0.0294***
	(0.0162)	(0.0053)	(0.0096)	(0.0070)	(0.0090)
After 2010 x Black or Hispanic	-0.0165***	0.0360***	0.0044*	0.0016	0.0044**
	(0.0045)	(0.0021)	(0.0025)	(0.0018)	(0.0020)
Constant	0.1493***	0.0032***	0.0353***	0.0163***	0.0160***
	(0.0061)	(0.0010)	(0.0032)	(0.0022)	(0.0021)
P-value: W = BH	0.2503	0.0000	0.3470	0.2539	0.0066
Observations	52,745	52,745	52,745	52,745	52,745

Notes. Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. 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."

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

Table A10b. Missing Mass in the Distribution of Drug Amounts by Race, with Various Time Trend Controls and State FEs

	Pr(< 280g)	Pr(280-290g)	Pr(> 290g)
	(1)	(2)	(3)
Panel A. No Interaction with Time Trend			
After 2010 x White	-0.0685***	0.0120**	0.0566***
	(0.0151)	(0.0055)	(0.0143)
After 2010 x Black or Hispanic	-0.0602***	0.0343***	0.0259***
	(0.0051)	(0.0023)	(0.0047)
Constant	0.9372***	0.0059***	0.0569***
	(0.0053)	(0.0013)	(0.0052)
P-value: W = BH	0.5840	0.0001	0.0330
Observations	52,678	52,678	52,678
Panel B. Interaction with Linear Time Trend			
After 2010 x White	-0.0403*	0.0164**	0.0240
	(0.0229)	(0.0083)	(0.0218)
After 2010 x Black or Hispanic	-0.0601***	0.0345***	0.0256***
	(0.0064)	(0.0033)	(0.0057)
Constant	0.9078***	0.0043**	0.0880***
	(0.0100)	(0.0020)	(0.0098)
P-value: W = BH	0.4063	0.0418	0.9418
Observations	52,678	52,678	52,678
Panel C. Interaction with Quadratic Time Trends			
After 2010 x White	0.0031	0.0133	-0.0164
	(0.0303)	(0.0099)	(0.0291)
After 2010 x Black or Hispanic	-0.0256***	0.0301***	-0.0045
	(0.0085)	(0.0040)	(0.0078)
Constant	0.8789***	0.0038	0.1173***
	(0.0192)	(0.0040)	(0.0188)
P-value: W = BH	0.3614	0.1150	0.6933
Observations	52,678	52,678	52,678

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about sample selection. 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 general model I estimate is:

$$(Charged\,X - Y g)_{it} = \alpha_0 + \beta_1(After2010 \times W)_{it} + \beta_2(After2010 \times BH)_{it} + \delta_1(After2010 \times W \times Trend)_{it} + \delta_2(After2010 \times BH \times Trend)_{it} + \gamma_1 BH + \phi_1(BH \times Trend) + Z_i + g(t)_t + \epsilon_{it}$$

$Trend$ takes on the value of zero (i.e. no trend interaction), a linear trend, or a quadratic trend. $g(t)_t$ is a linear trend when no trend interactions are used and when the linear trend interaction is used. $g(t)_t$ is a quadratic trend when the quadratic trend interactions are used. Figures A7j-k show the total share of cases below 280g and above 280g over time, by race. For these shares, there are considerable trends over time, especially for white offenders. To quantify the break in those trends after 2010, I estimate case-level regressions that interact the dummy variable for after 2010 with a linear time trend centered at zero in 2011. Panel (a) shows the estimates without accounting for these time trends, and as a result, column 3 indicates that white offenders are more likely to be charged with amounts greater than 290g after 2010, relative to black and Hispanic offenders. This is true, but it is due to a substantial rise in cases above 290g for white offenders that begins in 2005. Panels (b) and (c) account for this by estimating the break in the trend after 2010. Both panels indicate that white, black, and Hispanic offenders have similar (and small) trend breaks in their share of cases above 290g. Likewise, both panels show bunching at 280-290g, a racial disparity in bunching, and evidence that the excess mass at 280-290g is drawn from cases that would have been charged below 280g prior to 2010. All specifications include state fixed-effects (Z_i).
*** p<0.01, ** p<0.05, * p<0.1

Table A10c. Missing Mass in the Distribution of Drug Amounts, Post-2007 Only

Panel A. Analysis of Changes in the 0-100g Range.

	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	0.0246***	-0.0710***	0.0323***	-0.0098***	-0.0120***
	(0.0042)	(0.0055)	(0.0044)	(0.0033)	(0.0042)
Constant	0.1065***	0.2920***	0.1134***	0.0751***	0.1263***
	(0.0024)	(0.0035)	(0.0025)	(0.0021)	(0.0026)
Predicted Change from Conceptual Model	Increase	Decrease	Increase	Decrease	Decrease
Observations	25,893	25,893	25,893	25,893	25,893

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)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	-0.0108**	0.0336***	0.0050*	0.0026	0.0056**
	(0.0050)	(0.0021)	(0.0027)	(0.0019)	(0.0022)
Constant	0.1886***	0.0062***	0.0443***	0.0207***	0.0269***
	(0.0031)	(0.0006)	(0.0016)	(0.0011)	(0.0013)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	25,893	25,893	25,893	25,893	25,893

Notes. Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams and sentenced from 2007-2015. The predicted change from the conceptual model of prosecutor behavior in Section II.B is displayed in the row labeled "predicted change from conceptual model."
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A10d. Missing Mass in the Distribution of Drug Amounts, Trial Cases Only

Panel A. Analysis of Changes in the 0-100g Range.

	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	0.0294	-0.0530**	0.0248	-0.0120	-0.0591***
	(0.0181)	(0.0216)	(0.0171)	(0.0137)	(0.0126)
Constant	0.1104***	0.2592***	0.0984***	0.0831***	0.1112***
	(0.0064)	(0.0089)	(0.0061)	(0.0056)	(0.0064)
Predicted Change from Conceptual Model	Increase	Decrease	Increase	Decrease	Decrease
Observations	2,841	2,841	2,841	2,841	2,841
R-squared	0.030	0.020	0.006	0.008	0.007

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)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	-0.0392**	0.0749***	0.0030	0.0217*	0.0085
	(0.0199)	(0.0131)	(0.0124)	(0.0111)	(0.0111)
Constant	0.2050***	0.0033***	0.0562***	0.0281***	0.0389***
	(0.0082)	(0.0012)	(0.0047)	(0.0034)	(0.0039)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	2,841	2,841	2,841	2,841	2,841
R-squared	0.012	0.022	0.007	0.006	0.010

Notes. Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams and cases that end in a jury trial. The predicted change from the conceptual model of prosecutor behavior in Section II.B is displayed in the row labeled "predicted change from conceptual model."
*** p<0.01, ** p<0.05, * p<0.1

Table A11. Sentencing Consequences of Being Above the Threshold Amount

	Years Sentenced					
	(1)	(2)	(3)	(4)	(5)	(6)
Above 280g	-0.580**	0.0621			0.00410	-0.0576
	(0.289)	(0.691)			(0.294)	(0.461)
Above 280g x After 2010	2.332***	2.181**			0.971*	2.836***
	(0.508)	(1.102)			(0.535)	(0.842)
Above 50g			0.755***	0.955***	1.469***	2.101***
			(0.128)	(0.158)	(0.180)	(0.227)
Above 50g x After 2010			-1.387***	-1.063***	-1.298***	-2.058***
			(0.270)	(0.357)	(0.451)	(0.445)
Constant	12.93***	11.48***	9.664***	9.540***	13.12***	14.08***
	(0.170)	(0.565)	(0.114)	(0.116)	(3.298)	(3.709)
Bandwidth	±250g	±50g	±250g	±50g	±250g	±250g
Includes Life & <1 Month	No	No	No	No	No	Yes
Observations	29,767	2,800	49,154	14,713	29,064	31,134
R-squared	0.037	0.015	0.070	0.035	0.038	0.031

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. The coefficients in columns 1-2 are estimated from the following regression discontinuity style model:

$$Sentence_{it} = \alpha + \beta_1 Above280_{it} + \beta_2 Amount_{it} + \beta_3 (Above280 \times Amount)_{it} + \delta_1 (Above280 \times After2010)_{it} \\ + \delta_2 (Amount \times After2010)_{it} + \delta_3 (Above280 \times Amount \times After2010)_{it} + g(t)_t + \epsilon_{it}$$

where $Amount_{it}$, the running variable, is the amount of drugs centered at the 280g mandatory minimum, $After2010_{it}$ is a dummy variable equal to one if the case is sentenced after 2010, and $Above280_{it}$ is a dummy variable equal to one if the case involves 280g or more of crack-cocaine. Columns 3-4 estimate equation (4) around the 50g threshold instead of the 280g threshold. Columns 5-6 estimate the sentencing penalty around the 50g threshold and the 280g threshold simultaneously. In addition, all specifications above include a time trend to capture the gradual decline in sentences over time. Column 6 includes life sentences (coded as 70 years) and sentences less than 1 month (coded as 0 years). I do not find significant differences in these sentencing discontinuities by race. I include the R-squared in this table because the dependent variable is continuous. Figures 3a-d show graphical evidence of the sentencing penalty. Figure A9 shows that the estimate of the sentencing penalty from model (5) is robust to many different bandwidths from 10g to 250g. *** p<0.01, ** p<0.05, * p<0.1

Table A12. Bunching Analysis for Potential Mechanisms, Alternative Results

Panel A. Analysis of Bunching in State Convictions and in Drug Seizures

	Pr(200-400g) (1)	Pr(200-400g) (2)	Pr(280-290g) (3)	Pr(280-290g) (4)
After 2010	0.00358		0.0185	
	(0.00873)		(0.0444)	
After 2010 x White		0.0068		-0.0008
		(0.0116)		(0.0554)
After 2010 x Black or Hispanic		0.0017		0.0192
		(0.0095)		(0.0488)
Constant	0.103***	0.1018***	0.2132***	0.1615***
	(0.00616)	(0.0068)	(0.0297)	(0.0379)
Data Analyzed	FL Convictions	FL Convictions	NC Convictions	NC Convictions
Drugs Included	Cocaine, all types, Weight Only	Cocaine, all types, Weight Only	Cocaine, all types	Cocaine, all types
P-value: W = BH	-	0.6484	-	0.2382
Observations	12,194	12,194	843	843

Panel B. Analysis of Bunching in Drug Seizures and Final Sentencing

	Pr(280-290g) (6)	Pr(200-400g) (7)	Pr(200-400g) (8)	Pr(280-290g) (9)
After 2010	-0.000186**		0.0332**	
	(8.67e-05)		(0.0162)	
After 2010 x White		0.0002		0.0038
		(0.0002)		(0.0513)
After 2010 x Black or Hispanic		-0.0003***		0.0346**
		(0.0001)		(0.0164)
Constant	0.000422***	0.0003***	0.143***	0.1558***
	(4.94e-05)	(0.0001)	(0.0120)	(0.0219)
Data Analyzed	NIBRS, Full Coverage States	NIBRS, Full Coverage States	USSC Sentencing, NC only	USSC Sentencing, NC only
Drugs Included	Crack-cocaine	Crack-cocaine	Cocaine, all types	Crack-cocaine
P-value: W = BH	-	0.0830	-	0.5469
Observations	219,515	219,515	4,376	4,376

Notes. Robust standard errors in parentheses. When possible, the specifications above use a sample of offenses with drug amounts between 0 grams and 1000 grams. Analyses of state-level drug convictions do not make this restriction since the state reports broad drug weight categories instead of specific amounts. When broad categories (200-400g) are analyzed, a linear trend in year is included. The row "P-value: W= BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." In Panel A: columns 1-2 show an analysis of reported drug amounts for state-level drug convictions in Florida that restricts to cases where some weight range is listed in the offense description, columns 3-4 show an analysis of state-level drug convictions in North Carolina (a state where only some offenses specify the type of drug involved). Columns 5-6 show an analysis of weights for seized drugs reported to the FBI through the National Incident Based Reporting System (limiting to states that have full coverage from 2012-2015 and have at least 90% coverage from 2008-2015), Finally, columns 7-8 show an analysis of weights from USSC sentencing data for federal convictions in NC using broad drug categories and all types of cocaine.

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

Table A13. Variation in Bunching at 280-290g By Type of Agency Sending the Case

	280-290g (1)	280-290g (2)	280-290g (3)	Weight (g) (4)
After 2010	0.0826***	0.0760***	0.0989***	26.09***
	(0.0180)	(0.0191)	(0.0129)	(5.659)
After 2010 × Any Federal	-0.00889			
	(0.0190)			
After 2010 × Only Federal		-0.00263		
		(0.0202)		
After 2010 × FBI			0.0160	52.99***
			(0.0198)	(11.29)
After 2010 × ATF			-0.0732***	-15.03**
			(0.0143)	(6.953)
After 2010 × State/local			-0.0229	-7.648
			(0.0231)	(11.45)
After 2010 × DEA & State/local			-0.0133	-3.980
			(0.0383)	(19.46)
After 2010 × Joint state/local			0.0148	7.345
			(0.0507)	(25.60)
After 2010 × ATF & State/local			-0.00860	-9.386
			(0.0388)	(13.18)
After 2010 × FBI & State/local			-0.0619	-17.32
			(0.0386)	(22.44)
Constant	0.00342***	0.00360***	0.00481***	77.73***
	(0.00121)	(0.00136)	(0.000876)	(1.523)
Observations	17,042	15,016	17,042	17,042

Notes. Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. Column 1 interacts the after 2010 dummy variable with a dummy variable equal to one when the agency recorded as sending the case involves a federal agency (i.e. DEA, ATF, FBI). This includes agencies recorded as a federal agency joint with a state/local task force. Column 2 interacts the after 2010 variable with a variable equal to one when the agency sending the case is strictly federal (i.e. not including any involvement from state/local authorities). Column 2 does not include "joint" investigations in the sample. Column 3 provides more detail by interacting the after 2010 dummy variable with dummy variables for the top agencies (with the DEA as the reference category). Most agencies have similar levels of bunching at 280-290g post-2010. Two agencies have considerably lower levels, but as column 4 shows, those agencies are involved with lower drug weight cases, in general.

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

Table A14. Offender Drug-Holding Behavior by Race, After Fair Sentencing Act in 2010, Full Coverage States

	Weight	Pr(280-290g)	Weight	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-280g)	Pr(270-280g)	Pr(280-290g)	Pr(>290g)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After 2010 x White			-0.6018	0.0302***	-0.0210***	-0.0033**	-0.0058***	-0.0001	0.0001	-0.0002
			(0.5999)	(0.0041)	(0.0037)	(0.0016)	(0.0013)	(0.0000)	(0.0002)	(0.0007)
After 2010 x Black			-2.8015***	0.0403***	-0.0172***	-0.0064***	-0.0143***	-0.0001***	-0.0002**	-0.0020***
			(0.2504)	(0.0027)	(0.0025)	(0.0011)	(0.0009)	(0.0000)	(0.0001)	(0.0003)
Black	2.503***	9.21e-05	3.0414***	-0.1125***	0.0825***	0.0137***	0.0148***	0.0001	0.0002	0.0013***
	(0.260)	(0.000102)	(0.2885)	(0.0025)	(0.0023)	(0.0010)	(0.0009)	(0.0001)	(0.0001)	(0.0004)
Constant	10.01***	0.000454***	9.7586***	0.7503***	0.1856***	0.0310***	0.0284***	0.0002**	0.0004***	0.0043***
	(0.426)	(0.000152)	(0.4417)	(0.0040)	(0.0036)	(0.0016)	(0.0014)	(0.0001)	(0.0002)	(0.0006)
Observations	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043
P-value: W = B	-	-	0.0007	0.0408	0.3969	0.1075	0.0000	0.3308	0.1266	0.0205

Notes. Robust standard errors in parentheses. This analysis uses the weights of seized drugs reported to the FBI through the National Incident Based Reporting System. Ethnicity is not consistently recorded in NIBRS over this time period. As such, I refer to offenders as black or white, omitting the Hispanic label used in previous analyses. Columns 1-3 show the relationship between race of offender and drug weight seized, in general. Column 4 shows how the weight of an offender's seized drugs changes by race after 2010. Columns 5-11 show how the probability an offender's seized drugs are in a certain bin changes by race after 2010. All specifications include state fixed effects and controls for age and sex. The row "P-value: W= B" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black." The sample is limited to states that have full coverage from 2012-2015 and have at least 90% coverage from 2008-2015.

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

Table A15. Relationship between Bunching in EOUSA and Imputed Defendant Race

	280-290g, Missing = 0	280-290g	280-290g, Missing = 0	280-290g	280-290g, Missing = 0
	(1)	(2)	(3)	(4)	(5)
After 2010	0.0241***	-0.0318	-0.0153**	-0.00536	-0.00511
	(0.00180)	(0.0196)	(0.00654)	(0.0229)	(0.00826)
After 2010 × % Black or Hispanic		0.123***	0.0457***	0.0793***	0.0303***
(for Cases Sentenced in District-Month)		(0.0295)	(0.01000)	(0.0282)	(0.00984)
Constant	0.00159***	-0.00193	-0.00111	-0.00202	-0.000842
	(0.000195)	(0.00319)	(0.00130)	(0.00633)	(0.00259)
Prosecutor FEs	NO	NO	NO	YES	YES
Observations	49,342	13,384	32,751	13,384	32,751

Notes. Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. Column 1 displays the main bunching result using a dependent variable that is equal to one when the drug weight in the case is between 280-290g and is equal to zero if it is not in that range. Importantly, "280-290g, Missing=0" is also coded as zero if the drug weight field is missing. This is especially relevant for cross-district analyses because weight missingness varies substantially across districts. Coefficients are estimated from the following regression for column 1:

$$(Charged\ 280-290g,\ Missing=0)_{it} = \alpha_0 + \beta_1 After2010_{it} + \epsilon_{it}$$

Columns 2-5 interact the after 2010 dummy variable with a probabilistic estimate of defendant race (race is not available in the EOUSA files). To impute defendant race, I match EOUSA information about sentence year-month to USSC information about the racial composition of sentences in each sentence year-month. I code "% Black or Hispanic" equal to the fraction of offenders sentenced in a year-month who are black or Hispanic. In columns 4-5, I include prosecutor fixed effects. Specifications with the race and after 2010 interactions also include a variable equal to % black and Hispanic offenders in the district-month. The number of observations falls because not all cases that enter EOUSA end in a sentence. Coefficients are estimated from the following regression for columns 2 and 3 (with only the dependent variable changing):

$$\begin{aligned}(Charged\ 280-290g)_{it} = \alpha_0 + \beta_1(After2010)_{it} + \\ \beta_2(After2010 \times \%BlackOrHispanic)_{it} + \%BlackOrHispanic_{it} + \epsilon_{it}\end{aligned}$$

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

Table A16. Relationship between Bunching in EOUSA and State-level Racial Animus

	280-290g (1)	280-290g, Missing = 0 (2)	# of Attys in State who Bunch at 280g
After 2010	0.0756***	0.0163***	-
	(0.0123)	(0.00287)	-
Above Med. Racial Animus	-0.00187	-0.000390	1.737**
	(0.00122)	(0.000447)	(0.690)
After '10 × Above Med. Racial Animus	0.00150	0.0106***	-
	(0.0138)	(0.00365)	-
Constant	0.00520***	0.00182***	-
	(0.00111)	(0.000388)	-
Observations	19,241	49,051	51

Notes. Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. See Table A15 for a discussion of the "280-290, Missing=0" dependent variable. Columns 1 and 2 interact the after 2010 dummy variable with a dummy variable equal to one when the state where the case is received is above the median level of racial animus and equal to zero if it is below the median level. Coefficients are estimated from the following regression for columns 1 and 2 (with only the dependent variable changing):

$$(Charged\,280-290g)_{it} = \alpha_0 + \beta_1(After2010)_{it} + \beta_2(After2010 \times AboveMedRA)_{it} + AboveMedRA_{it} + \epsilon_{it}$$

Since racial animus is a measure that varies across districts, column 2 results are particularly noteworthy (using the "missing included" version of 280-290g accounts for some of the cross-district variation in drug weight reporting). Finally, column 3 estimates a state-level regression of the number of bunching attorneys in the state (defined as an attorney whose fraction of cases at 280-290g post-2010 is above the average fraction at 280-290g pre-2010) on the above median racial animus dummy variable.

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

Table A17. Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors

	Atty. with 5+ Cases			Atty. with 15+ Cases		
Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010						
	Below 280g	280-290g	Above 290g	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)	(4)	(5)	(6)
Atty. Bunches at 280-290g Post-2010	-0.2193***	0.2421***	-0.0228	-0.1143	0.1882***	-0.0739
(15+ cases post-2010)	(0.0459)	(0.0339)	(0.0272)	(0.0806)	(0.0447)	(0.0640)
Constant	0.9309***	-	0.0691***	0.8855***	-	0.1145*
	(0.0242)	-	(0.0242)	(0.0617)	-	(0.0617)
Observations	1,647	1,647	1,647	699	699	699
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010						
	Below 280g	280-290g	Above 290g	Below 280g	280-290g	Above 290g
	(7)	(8)	(9)	(10)	(11)	(12)
Atty. Bunches at 50-60g Pre-2010	-0.0665***	0.0467***	0.0198	-0.0863***	0.0611***	0.0252
(15+ cases pre-2010)	(0.0245)	(0.0169)	(0.0151)	(0.0263)	(0.0167)	(0.0178)
Constant	0.9258***	0.0335***	0.0407***	0.9466***	0.0153	0.0382***
	(0.0168)	(0.0111)	(0.0115)	(0.0172)	(0.0096)	(0.0139)
Observations	1,278	1,278	1,278	956	956	956

Notes. Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(Charged\,X - Y g)_i = \alpha_0 + \beta_1 AttyBunchesAt280g_i + \epsilon_i$$

where $AttyBunchesAt280g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). These regressions are restricted to post-2010 cases and to prosecutors with 5+ cases post-2010 in columns 1-3 and with 15+ cases post-2010 in columns 4-6. Note, column (2) is a mechanical relationship, hence the missing standard error. Coefficients in panel B are estimated from the following regression for each range:

$$(Charged\,X - Y g)_i = \alpha_0 + \beta_1 AttyBunchesAt50g_i + \epsilon_i$$

where $AttyBunchesAt50g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 50g definition (i.e. the fraction of their cases that are from 50-60g is above the average fraction of 50-60g cases post-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 50-60g is at or below the average fraction of 50-60g cases post-2010). These regressions are restricted to post-2010 cases and to prosecutors with 5+ cases pre-2010 in columns 7-9 and with 15+ cases pre-2010 in columns 10-12.
*** p<0.01, ** p<0.05, * p<0.1

Table A18. Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors, Leave-One-Out Classification

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	-0.114*	0.149***	-0.0354
(Leaving out current case in calculation)	(0.0659)	(0.0435)	(0.0463)
Constant	0.891***	0.0272***	0.0816*
	(0.0432)	(0.00765)	(0.0436)
Observations	971	971	971
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Pct. of Cases Bunched at 280-290g	-0.505***	0.527***	-0.0227
(Leaving out current case in calculation)	(0.116)	(0.0717)	(0.0976)
Constant	0.891***	0.0380***	0.0708**
	(0.0346)	(0.00791)	(0.0349)
Observations	971	971	971

Notes. Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(Charged\,X - Y\,g)_i = \alpha_0 + \beta_1 AttyBunchesAt280g_i + \epsilon_i$$

where $AttyBunchesAt280g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). **The classification for each bunching attorney is based on all cases excluding the current observation (i.e. a leave-one-out procedure).** Coefficients in panel B are estimated from the following regression for each range:

$$(Charged\,X - Y\,g)_i = \alpha_0 + \beta_1 PctBunching280g_i + \epsilon_i$$

where $PctBunchingAt280g$ is equal to the prosecutor's fraction of cases at 280-290g post-2010 (excluding the current observation) minus the average fraction of cases at 280-290g pre-2010. These regressions are restricted to post-2010 cases and to prosecutors with 10+ cases post-2010.

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

Table A19. Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors, with Bootstrapped SEs

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	-0.1794***	0.2170***	-0.0376
	(0.0659)	(0.0371)	(0.0510)
Constant	0.9184***	-	0.0816*
	(0.0435)	-	(0.0435)
Observations	989	989	989
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Atty. Bunches at 50-60g Pre-2010	-0.0785***	0.0575***	0.0211
	(0.0299)	(0.0177)	(0.0180)
Constant	0.9359***	0.0233**	0.0408***
	(0.0170)	(0.0105)	(0.0133)
Observations	1,135	1,135	1,135

Notes. Standard errors are calculated from 25 replications of a bootstrapping procedure that samples cases (with replacement) clustered at the prosecutor-level and calculated the bunching dummy variables within each sample. The standard errors for the constant terms are not calculated in this way; robust errors clustered at the prosecutor-level are used. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(Charged\,X - Y g)_i = \alpha_0 + \beta_1 AttyBunchesAt280g_i + \epsilon_i$$

where $AttyBunchesAt280g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). These regressions are restricted to post-2010 cases (for columns 1-3) and to prosecutors with 10+ cases post-2010. Note, column (2) is a mechanical relationship, hence the missing standard error. Coefficients in panel B are estimated from the following regression for each range:

$$(Charged\,X - Y g)_i = \alpha_0 + \beta_1 AttyBunchesAt50g_i + \epsilon_i$$

where $AttyBunchesAt50g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 50g definition (i.e. the fraction of their cases that are from 50-60g is above the average fraction of 50-60g cases post-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 50-60g is at or below the average fraction of 50-60g cases post-2010). These regressions are restricted to post-2010 cases (for columns 5-8) and to prosecutors with 10+ cases pre-2010.
*** p<0.01, ** p<0.05, * p<0.1

Table A20. Persistence of Attorney-level Bunching Across Districts, from Analysis of Movers

	Pr(Atty. Bunches at 10-Year Mandatory Minimum in 2nd District)			
	(1)	(2)	(3)	(4)
Atty. Bunches at 10-Year MM in 1st District	0.184*	0.162**	0.263**	0.154*
	(0.0936)	(0.0816)	(0.108)	(0.0829)
Constant	0.500***	0.432***	0.462***	0.440***
	(0.0700)	(0.0580)	(0.0809)	(0.0577)
Bunching classification	280-290g, National	280-290g, Missing=0, National	280-290g, District	280-290g, Missing=0, District
Observations	109	148	79	144

Notes. Robust standard errors are in parentheses. The estimates in this table are based on the EOUSA data. For this analysis, I identify the attorneys who switch districts at some point in their career (using their initials recorded in the EOUSA case management system). I then identify the set of those attorneys who bunch at a 10-year mandatory minimum in their first district. I also limit the sample to attorneys who have at least 5+ cases in their first district and 5+ cases in their second district (this maintains the 10+ restriction but spreads it evenly across districts). Since I am analyzing movers, it is almost always the case that the cases in their first district are pre-2010 cases, meaning that the bunching classification is determined based on bunching at 50-60g. Finally, I regress an indicator equal to one if the attorney bunches at the 10-year threshold in their second district on whether they bunched at the 10-year threshold in their first district. I do this for four methods of classifying bunching attorneys. Columns 1 and 2 are detailed in Table A15. Columns 3 and 4 mirror those two approaches but define the "baseline" bunching at the district-level. For example, an attorney **i** in district **A** is defined as bunching at 50-60g in column 3 if their fraction of cases at 50-60g pre-2010 is above the fraction of cases at 50-60g in district **A** post-2010. In all cases, I find that an attorney who bunches above the mandatory minimum threshold in their first district is more likely to do so in their second district than an attorney who does not bunch above the mandatory minimum threshold in their first district.
*** p<0.01, ** p<0.05, * p<0.1

Table A21. Relationship between Various Bunching Ranges, Attorneys

	28-29g	28-29g	50-60g	280-290g	280-290g	280-290g
	(1)	(2)	(3)	(4)	(5)	(6)
Atty. Bunches at 280-290g Post-2010	0.144**	0.140**	0.182***			
	(0.0625)	(0.0590)	(0.0664)			
Atty. Bunches at 28-29g Post-2010				0.155***	0.0876**	
				(0.0544)	(0.0340)	
Atty. Bunches at 50-60g Pre-2010						0.0575***
						(0.0172)
Constant	0.131***	0.120***	0.155***	0.0826***	0.0479***	0.0233**
	(0.0241)	(0.0232)	(0.0288)	(0.0271)	(0.0149)	(0.0105)
Sample Years	2011-2017	2011-2017	2000-2010	2011-2017	2011-2017	2011-2017
Sample Restriction	0-280g	0-280g, 290-1000g	0-1000g	29-1000g	0-28g, 29-1000g	0-1000g
Observations	843	910	1,976	483	840	1,135

Notes. Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Columns 1-3 estimate the likelihood an attorney who bunches at 280-290g (i.e. who has a fraction of cases at 280-290g post-2010 that is above the average fraction of 280-290g cases pre-2010) also bunches at 28-29g post-2010, 28-29g post-2010, and 50-60g pre-2010, respectively. Column 1 limits the sample to cases with below 280g to avoid a mechanical relationship. Column 2 does this by excluding only the 280-290g range from the sample. Both approaches yield similar results. Column 3, since the dependent variable is based on pre-2010 data, uses the full range of cases (0-1000g). Columns 4-6 estimate the likelihood an attorney who bunches at 28-29g post-2010 or 50-60g pre-2010 also bunches at 280-290g post-2010. As before, columns 4 and 5 exclude the 28-29g range to avoid a mechanical relationship. 28-29g is relevant post-2010 because 28g is the threshold for the 5-year mandatory minimum after 2010. 50-60g is relevant pre-2010 because 50g is the threshold for the 10-year mandatory minimum prior to 2010. All regressions in this table use the sample of attorneys who have 10+ cases (post-2010 for columns 1-5; pre-2010 for column 6). In all cases, an attorney who bunches at one mandatory minimum threshold is more likely to bunch at a separate mandatory minimum threshold.

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

Table A22. Bunching at 280-290g and Drug Conspiracy Charges

	Pr(Lead Charge = Conspiracy)		
	(1)	(2)	(3)
Case recorded at 280-290g	0.396***	0.307***	0.249***
	(0.0326)	(0.0329)	(0.0361)
Constant	0.166***	0.255***	0.314***
	(0.00279)	(0.00487)	(0.0156)
Sample restriction	0-1000g	50-1000g	280-1000g
Observations	18,062	8,236	1,116

Notes. Robust standard errors are in parentheses. The estimates in this table are based on the EOUSA data. The dependent variable is an indicator equal to one if the lead charge on the case is a drug conspiracy charge. Drug conspiracy charges are a tool that prosecutors can use to increase the weight involved in the offense because the total weight of the conspiracy is applied to each offender deemed involved in the conspiracy. The independent variable is whether the case involves 280-290g. Cases with 280-290g are substantially more likely to carry a lead conspiracy charge. This is true even when limiting to cases with 280-1000g only (see column 3).
*** p<0.01, ** p<0.05, * p<0.1

Table A23. Effect of Alleyne v. US, Accounting for Missing Values

	Pr(Case's Drug Weight is Missing)	Pr(Case is Charged with 280-290g, Missing = 0)
	(1)	(2)
After June 17th, 2011-2016	-0.0211	0.00438
	(0.0309)	(0.00869)
After June 17th, 2013	-0.0219	-0.0389*
	(0.0702)	(0.0223)
Constant	0.834***	0.0243
	(0.0690)	(0.0269)
Bandwidth	±150 days	±150 days
Observations	6,182	6,182

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

$$\begin{aligned} Y_{it} &= \alpha_0 + \beta_1 AfterJune17_{it} + \beta_2 DaysFrom_{it} + \beta_3 (AfterJune17 \times DaysFrom)_{it} \\ &+ \delta_1 (AfterJune17 \times Year2013)_{it} + \delta_2 (DaysFrom \times Year2013)_{it} \\ &+ \delta_3 (AfterJune17 \times DaysFrom \times Year2013)_{it} + D_{it} + \epsilon_{it} \end{aligned}$$

where $AfterJune17$ is a dummy variable equal to one for cases received after June 17th in each year, $DaysFrom$, the running variable, is the date the case was received centered at zero on June 17th, and $Year2013$ is equal to one for cases received in 2013 (the year *Alleyne* is decided). In addition, all specifications above include day-of-week fixed effects, D_{it}, for the day the case is received. In column 1, Y_{it} is equal to one if the observation has a missing drug weight and equal to zero otherwise. There is little effect of Alleyne on the likelihood an observation has missing drug weight. In column 2, Y_{it} is equal to one if the drug weight is equal to 280-290g or if the drug weight is missing and equal to zero otherwise. There is still a decrease in bunching after Alleyne when accounting for missing values.
*** p<0.01, ** p<0.05, * p<0.1

Table A24. Degree of Bunching Post-2010 by Race and District-level Caseload Characteristics

	Pr(280-290g)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After '10 x White (W)	0.0172**	0.0183*	0.0161**	0.0197*	0.0131	0.0219***	0.0113*	0.0137	0.0128
	(0.0082)	(0.0100)	(0.0080)	(0.0102)	(0.0088)	(0.0083)	(0.0061)	(0.0089)	(0.0083)
After '10 x Black or Hispanic (BH)	0.0424***	0.0477***	0.0344***	0.0536***	0.0302***	0.0388***	0.0407***	0.0368***	0.0379***
	(0.0094)	(0.0035)	(0.0028)	(0.0035)	(0.0027)	(0.0029)	(0.0035)	(0.0030)	(0.0033)
After '10 x W x Char.	-0.0147	-0.0088	-0.0055	-0.0122	0.0008	-0.0191*	0.0007	-0.0043	-0.0028
	(0.0098)	(0.0116)	(0.0107)	(0.0115)	(0.0108)	(0.0103)	(0.0104)	(0.0107)	(0.0104)
After '10 x BH x Char.	-0.0187	-0.0222***	0.0027	-0.0363***	0.0124***	-0.0072*	-0.0077*	-0.0011	-0.0031
	(0.0114)	(0.0043)	(0.0043)	(0.0042)	(0.0044)	(0.0042)	(0.0045)	(0.0044)	(0.0044)
Constant	0.0024*	0.0054***	0.0054***	0.0053***	0.0053***	0.0053***	0.0045***	0.0046***	0.0046***
	(0.0014)	(0.0010)	(0.0011)	(0.0010)	(0.0011)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Characteristic	District-by-Year Above Med. # of Cases per Attorney	District Above Med. % of Guilty Cases	District Above Med. % of Declined Cases	District Above Med. % of Plea Cases	District Above Med. % of Cases Dismissed for 'Weak Evidence'	District Above Med. % of Cases Dismissed for 'Resources'	District Above Med. % Of Cases with Retained Counsel (based on '99-'02)	District Above Med. % Of Cases with Appointed Counsel	District Above Med. % Of Cases with Public Defender Counsel
P-value: W = BH	0.0246	0.0057	0.0297	0.0017	0.0609	0.0536	0.0000	0.0139	0.0049
P-value: W+Char. = BH+Char.	0.0000	0.0113	0.0007	0.0872	0.0001	0.0000	0.0191	0.0001	0.0003
Observations	52,731	52,745	52,745	52,745	52,745	52,745	49,851	49,851	49,851

Notes. Robust standard errors in parentheses. "Characteristic" or "Char." represents a dummy variable that is an district or district-by-year characteristic. The specific characteristic of interest is noted in the "Characteristic" row. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. 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)." Column 1 interacts the after 2010 by race dummy variables with a district-by-year dummy variable indicating if the district received above the median number of cases (per attorney) in the year. Column 2 studies districts above/below the median for percent of cases that end in a guilty verdict, column 3 studies districts above/below the median for percent of cases declined, and column 4 studies districts above/below the median for percent of cases that end in plea deals. Columns 5 and 6 study districts above/below the median for percent of cases declined due to "weak evidence" or "lack of resources" (as coded in the EOUSA case files, codes not present for all cases). Columns 7-9 use the USSC data from 1999-2002 on type of defense counsel to examine heterogeneity by type of defense counsel used in the district. Places with different rates of retained, appointed, or public defender defense counsel from 1999-2002 nevertheless have similar bunching at 280g post-2010.

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

Table A25. Relationship between Bunching at 280g and Judge Characteristics

	Pr(280-290g) (1)	Pr(280-290g) (2)	Pr(280-290g) (3)
After 2010	0.0928***	0.0891***	0.1042***
	(0.0093)	(0.0209)	(0.0151)
After 2010 ×White Judge		0.0045	
		(0.0233)	
After 2010 ×Republican Judge			-0.0197
			(0.0191)
Constant	0.0040***	0.0059**	0.0049***
	(0.0007)	(0.0024)	(0.0014)
Observations	8,359	8,359	8,359

Notes. Standard errors clustered at the judge level in parentheses. The estimates in this table are based on the EOUSA data. I can match judge race and political party to approximately half of the cases in the EOUSA data. For data on judge characteristics, I use the file provided by Cohen and Yang (2019). I estimate whether bunching at 280g is related to judge race or judge political party. Column (1) shows that the level of bunching is similar for cases where I can match judge characteristics. Column (2) shows that judge race does not affect bunching at 280g. Column (3) shows that judge political party does not affect bunching at 280g.

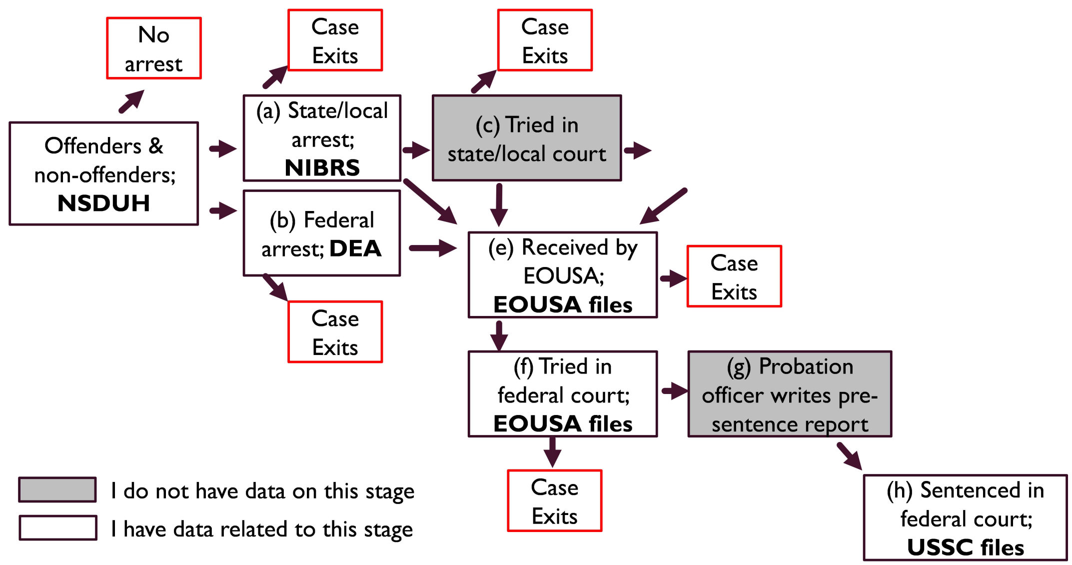
*** p<0.01, ** p<0.05, * p<0.1

Table A26. Relationship between Various Bunching Ranges, Judges

	28-29g	28-29g	50-60g	280-290g	280-290g	280-290g
	(1)	(2)	(3)	(4)	(5)	(6)
Judge Bunches at 280-290g Post-2010	-0.0129	-0.00857	0.0557			
	(0.0305)	(0.0286)	(0.0412)			
Judge Bunches at 28-29g Post-2010				-0.00207	-0.0144	
				(0.0523)	(0.0329)	
Judge Bunches at 50-60g Pre-2010						0.0175
						(0.0215)
Constant	0.155***	0.143***	0.199***	0.168***	0.108***	0.0723***
	(0.0195)	(0.0185)	(0.0243)	(0.0390)	(0.0250)	(0.0180)
Sample Restriction	0-280g	0-280g, 290-1000g	0-1000g	29-1000g	0-28g, 29-1000g	0-1000g
Observations	769	827	2,710	469	789	1,270

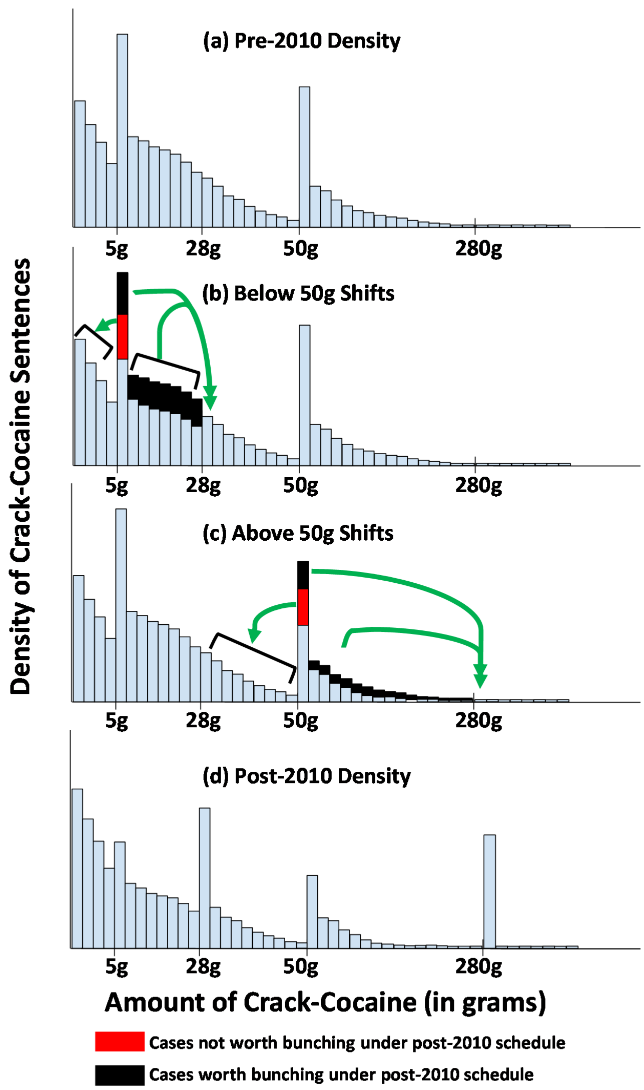
Notes. Standard errors clustered at the judge level in parentheses. The estimates in this table are based on the EOUSA data. See Table A21 for a discussion of the dependent and independent variables in column 1-6. The major difference is that these regressions examine judges classified as "bunching" at a given range. This is possible because the EOUSA files contain a judge ID for many cases. I use that judge ID to calculate the fraction of cases at 280-290g post-2010, 28-29g post-2010, and 50-60g pre-2010 for each judge. 28-29g is relevant post-2010 because 28g is the threshold for the 5-year mandatory minimum after 2010. 50-60g is relevant pre-2010 because 50g is the threshold for the 10-year mandatory minimum prior to 2010. All regressions in this table use the sample of judges who have 10+ cases (post-2010 for columns 1-5; pre-2010 for column 6). Judges who bunch at one mandatory minimum threshold are not more likely to bunch at other mandatory minimum thresholds.
*** p<0.01, ** p<0.05, * p<0.1

Figure A1. Graphical Illustration of Timeline from Arrest to Sentencing.



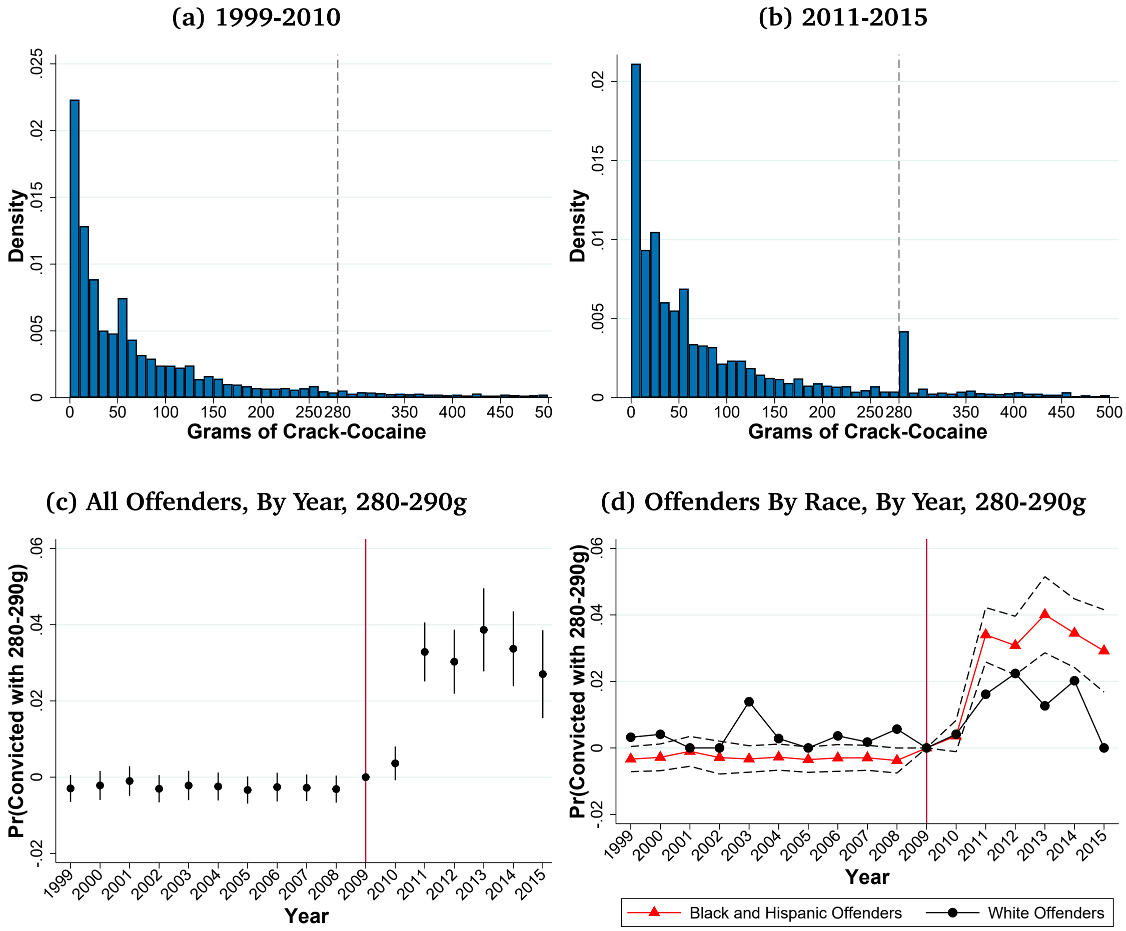
Notes. The figure above details the timeline from arrest to sentencing. Before arrest, the eventual arrestees come from the set of all people, some of whom are innocent and some of whom are guilty. Some individuals from this group are arrested by state/local police or federal police. Of those arrested by state/local police, their case can be dismissed, tried in state/local court, or passed on to federal authorities. Case tried in state/local court can leave the system if they are found not guilty, dismissed, etc., they can be convicted, or they can be sent to federal authorities. In fact, even convicted cases can be sent to federal authorities. Individuals arrested by federal police are typically referred to the EOUSA directly. Once a case is received by the EOUSA, it can leave the system via a dismissal, declination, etc., or it can be taken to federal court. For cases convicted in federal court, a probation officer prepares a pre-sentence report, and ultimately, the offender is sentenced. I have obtained data at nearly all of these steps. The two steps for which I lack data are in the middle of steps where bunching does not change, which suggests that nothing changes in the middle step.

Figure A2. Graphical Illustration of Conceptual Model, Prosecutor Responses to the FSA.



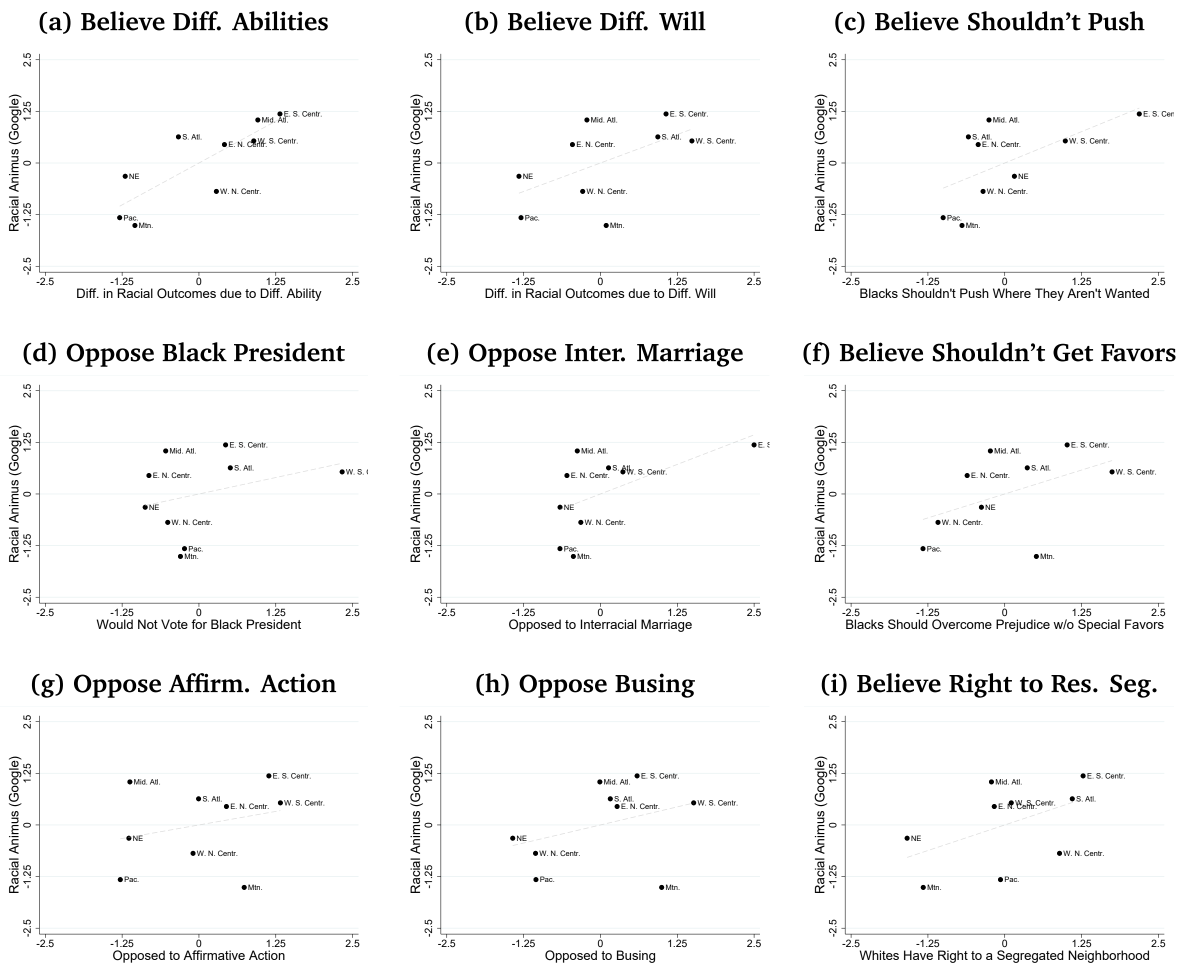
Notes. Panel (a) displays a hypothetical pre-2010 distribution of weights, with bunching at 5g and 50g due to round-number bias and prosecutor discretion. Panel (b) shows how the 0-5g, 5-28g, and 28-50g ranges will change after 2010. Some cases bunched at 5g will not be worth bunching at 28g (depicted in red), and they will shift into the 0-5g range. Some cases bunched at 5g and some cases from 5-28g will be worth bunching at 28g (depicted in black), and they will shift into the 28-50g range. Panel (c) illustrates a similar phenomena for the 50-280g range–some cases will shift down into the 28-50g range and some will shift up to the 280-290g range. Panel (d) shows the hypothetical post-2010 distribution of weights, with bunching at 5g and 50g due to round-number bias and bunching at 28g and 280g due to prosecutor discretion.

Figure A3. Changing Distribution of Drug Amounts Around 280g Pre- and Post-2010, USSC



Notes. Panels (a) and (b) plot the distribution of drug amounts recorded in federal crack-cocaine sentences starting at 0 grams and ending at 500 grams for 1999-2010 (when the mandatory minimum threshold was 50g) and 2011-2015 (when it was 280g). In panel (c), I estimate the main bunching coefficient by year (relative to 2010) and plot the coefficients with 90% confidence intervals. Panel (d) plots the coefficients and confidence interval for black and Hispanic offenders and the coefficients for white offenders (I do not include confidence intervals for white offenders because their estimates by year are extremely noisy).

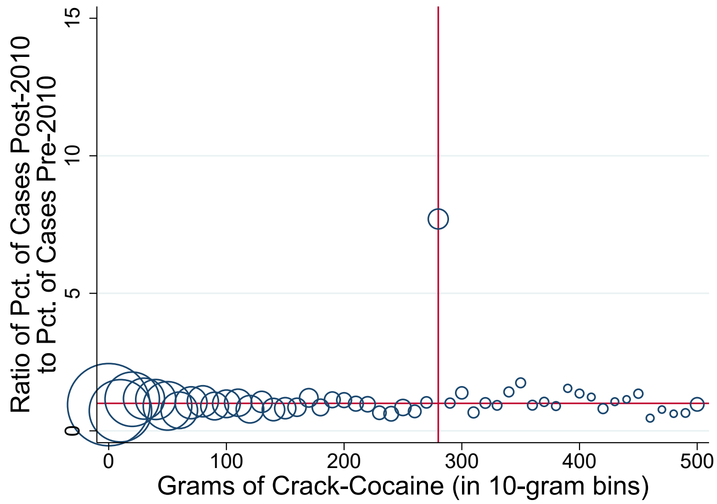
Figure A4. Relationship between Google Trends Racial Animus Measure and GSS Responses from Highly Educated Respondents on Attitudes about Race



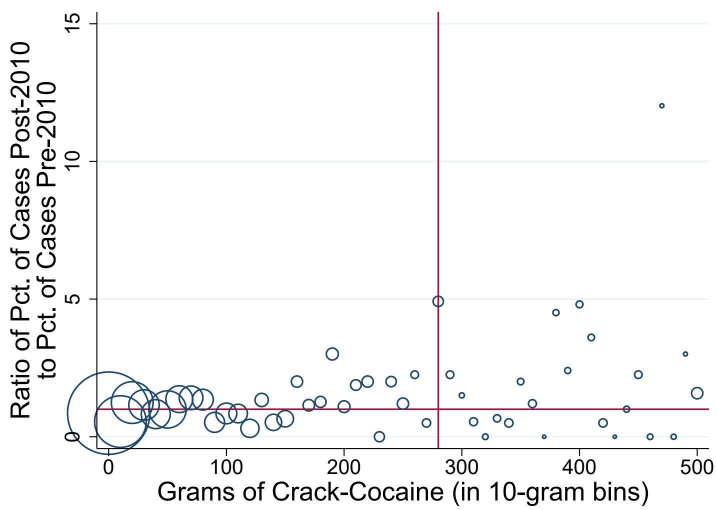
Notes. The figures above plot the relationship between the Google Trends racial animus measure (standardized and centered at zero) and various measures of attitudes about race from the General Social Survey (GSS) from 1972-2018 (not all questions are present in all years; also standardized and centered at zero). For the GSS measures, I limit the sample to respondents with a graduate degree or higher to test if the Google Trends racial animus measure is correlated with racial attitudes of highly educated people. The public sample of the GSS only includes region identifiers. I aggregate the Google Trends measure to the region level by taking the mean across all states in the region. The regions are: Northeast, West North Central, Pacific, Mountain, East North Central, Mid Atlantic, South Atlantic, West South Central, and East South Central. The GSS questions are: Do you believe... (a) racial differences in outcomes are due to different abilities by race (available 1977-2018), (b) racial differences in outcomes are due to different will by race (1977-2018), (c) black shouldn't push where they aren't wanted (1972-2002), (f) blacks should overcome prejudice without special favors (1994-2018), and (i) whites have a right to a segregated neighborhood (1972-1996)? And are you opposed to... (a) voting for a black president (1972-2010), (b) interracial marriage (1972-2002), (c) affirmative action (1994-2018), (d) desegregation busing (1972-1996)?

Figure A5. Bunching Ratio from 0-500g, USSC

(a) Bunching Ratio from 0-500g, Black and Hispanic Offenders



(b) Bunching Ratio from 0-500g, White Offenders

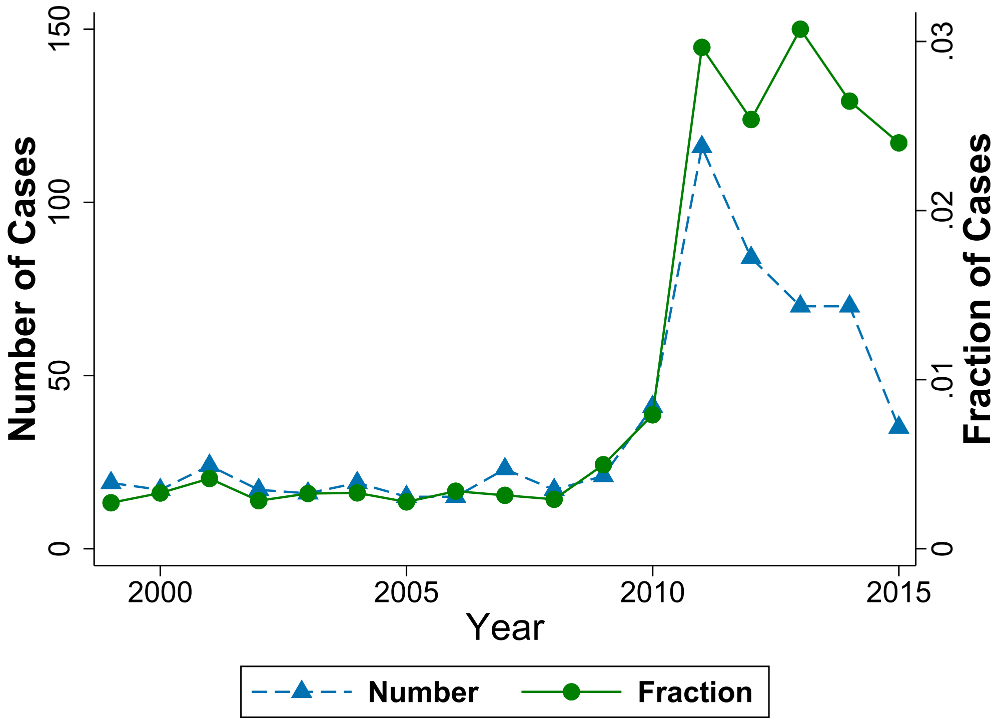


Notes. The figure above plots the bunching ratio for each 10-gram bin from 0-500 grams by race. The bunching ratio for each bin b is defined as follows:

$$\text{Bunching Ratio}_b = \frac{\text{\% of cases in b post-2010}}{\text{\% of cases in b pre-2010}}$$

If the distributions are the same pre- and post-2010, the bunching ratio will equal 1 (marked by the horizontal red line). If the ratio is above 1, there is a higher degree of bunching in bin b post-2010. If the ratio is below 1, there is a lower degree of bunching post-2010. The size of the marker for each bin b is weighted by the total number of cases in the bin pre- and post-2010 (relative to rest of the group included in the plot, not relative to the full sample).

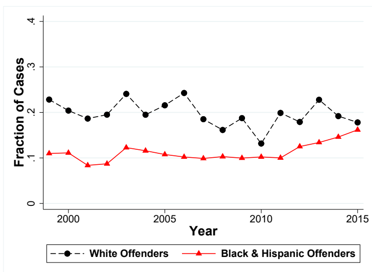
Figure A6. Number and Share of Offenses with 280-290g Over Time, USSC



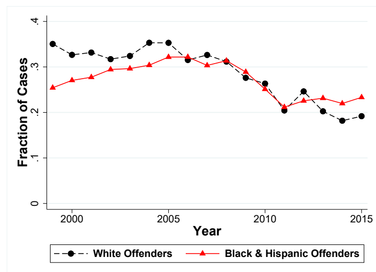
Notes. The figure above plots the total number of offenses with 280-290g over time and the share (or fraction) of cases with 280-290g over time.

Figure A7. Changing Distribution of Drug Weights Over Time, By Race, USSC

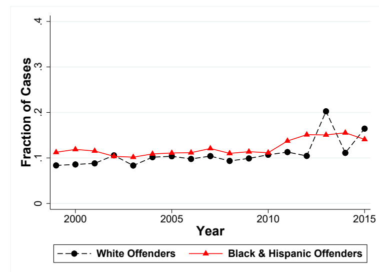
(a) 0-5 grams



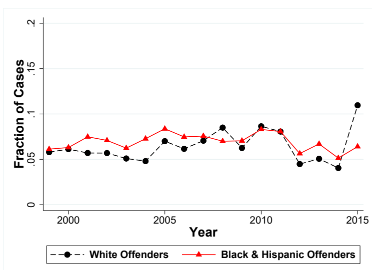
(b) 5-28 grams



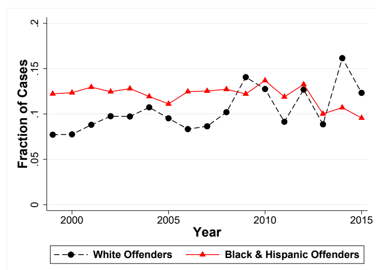
(c) 28-50 grams



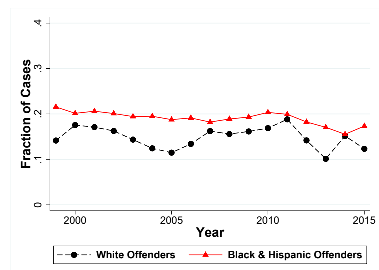
(d) 50-60 grams



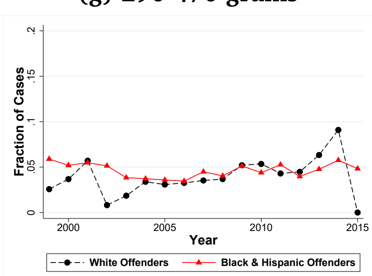
(e) 60-100 grams



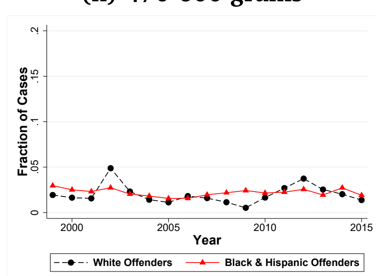
(f) 100-280 grams



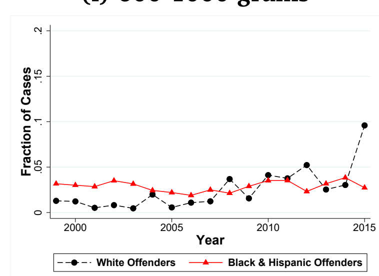
(g) 290-470 grams



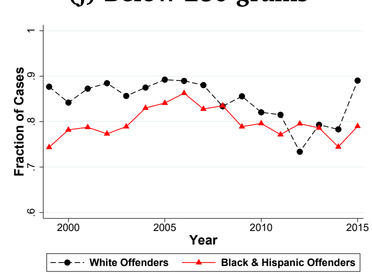
(h) 470-600 grams



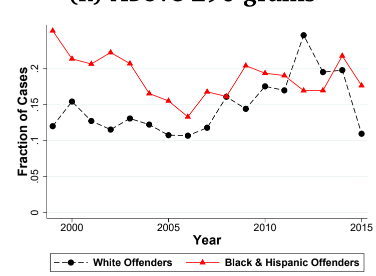
(i) 600-1000 grams



(j) Below 280 grams



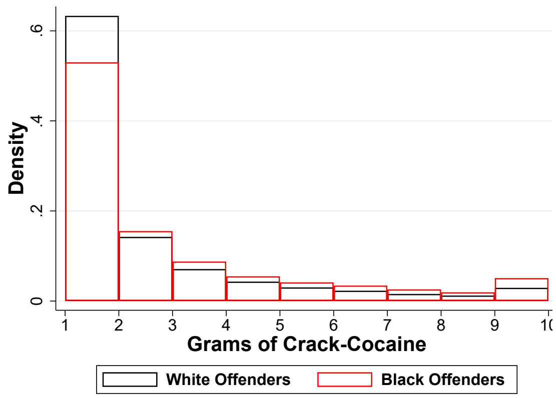
(k) Above 290 grams



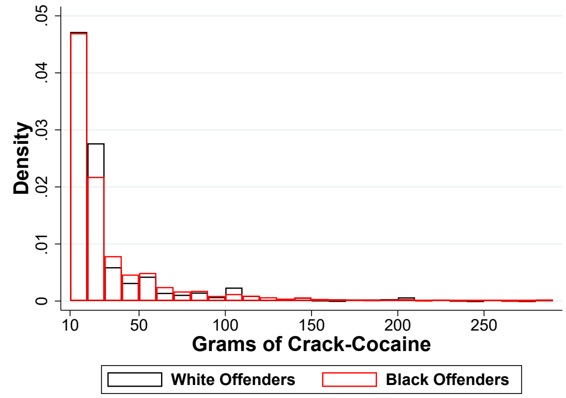
Notes. The figures above plot the share of cases in the specified range by year for white and black and Hispanic offenders. For example, panel (a) plots the share of cases with 0-5g (not including 5g) in each year from 1999-2015. Panel (b) plots the share of cases with 5-28g in each year from 1999-2015, and so on.

Figure A8. Alternative Figures for Conditional Racial Disparity Tests

(a) Drug Seizures by Race, 0-10g, NIBRS

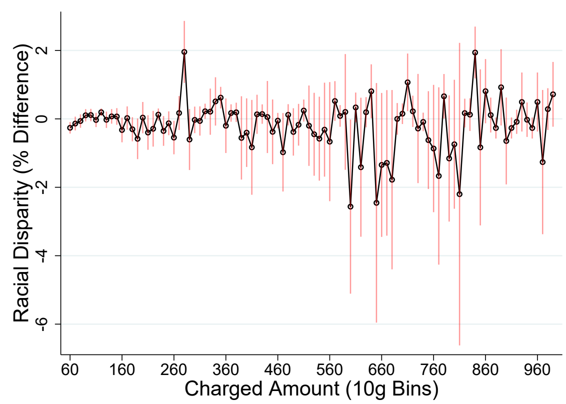


(a) Drug Seizures by Race, 10-280g, NIBRS



(c) Shifting from 60-500g by Race, USSC

(d) Shifting from 60-1000g by Race, USSC

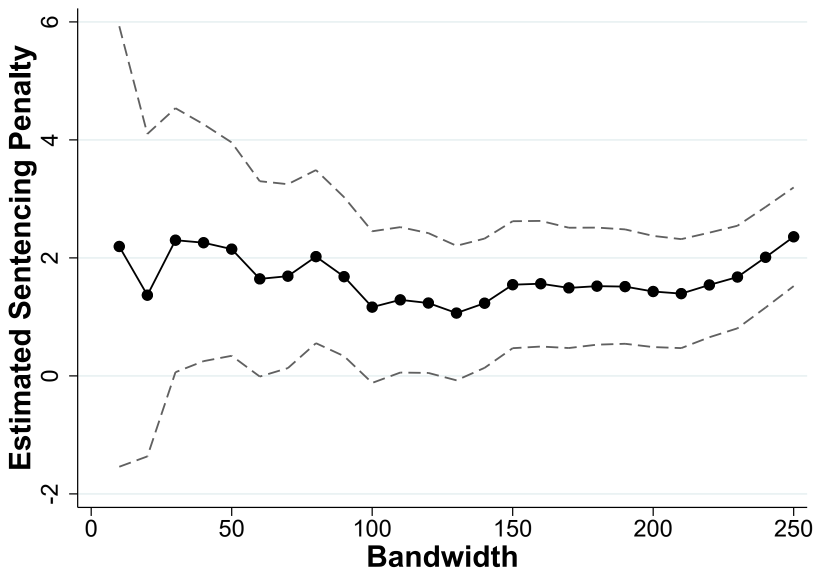


Notes. The figure in panel (a) plots the histograms of crack-cocaine amounts seized for white offenders and for black and Hispanic offenders from 0-10g. The white offenders are slightly over-represented at 1g, but otherwise, the distributions are very similar. The figure in panel (b) plots the histograms by race from 10-280g. White offenders are slightly over-represented at 20-30g, but otherwise, the distributions are very similar. These figures use the balanced sample of agencies (i.e. agencies that are present in all 16 years) in NIBRS. Panels (c) and (d) plot the coefficient δ^X for each 10g bin starting at X divided by the share of cases in that 10g bin (to calculate a percent difference).

$$(12) \quad (\text{Charged X-Yg})_{it} = \alpha + \delta^X(\text{After2010}\times\text{BlackOrHispanic})_{it} + \gamma\text{After2010}_{it} + \lambda\text{BlackOrHispanic}_i + \epsilon_{it}$$

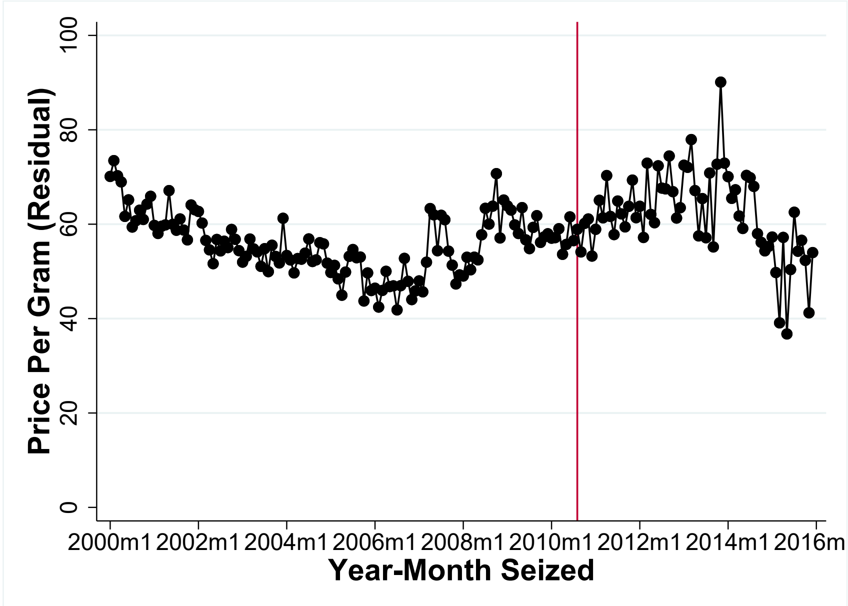
Since estimates are noisier at higher amounts, panel (c) shows the estimates for amounts from 0-500g alone and panel (d) shows the estimates for amounts from 0-1000g.

Figure A9. Sentencing Discontinuity Robust to Multiple Bandwidths



Notes. The figure above plots the sentencing penalty of crossing the 280g mandatory minimum threshold after 2010, as estimated using the RD difference-in-difference model specified in equation (5) of the main text. The dashed lines are 90% confidence intervals. Estimates using a quadratic in polynomial are similar in magnitude but slightly noisier. The bandwidths used in the figure above range from 10g to 250g, in 10g intervals.

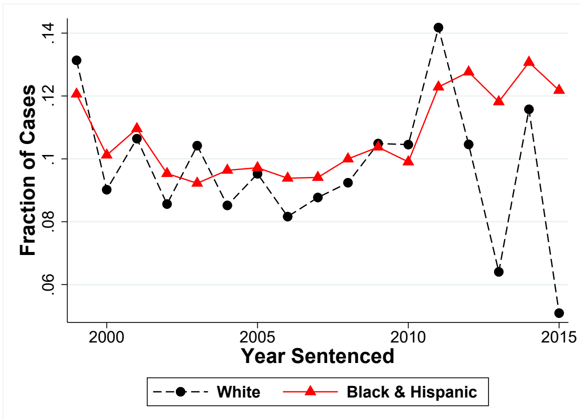
Figure A10. Drug Prices Before and After the Fair Sentencing Act, DEA



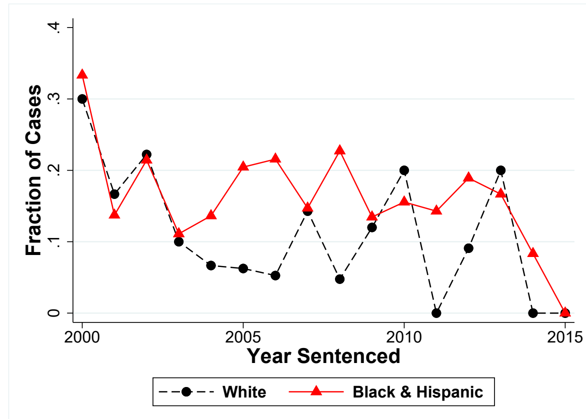
Notes. Panel (a) plots the drug price per gram (conditional on state, drug potency, type of drug, month seized, and a linear trend in year) against the year-month the drugs were seized. Outliers above the 95th percentile ($200 per gram) and below the 5th percentile ($20 per gram) are excluded. The price is smooth and increasing through the date the Fair Sentencing Act was implemented. In other words, there is no clear price response in the illegal drug market, at least in the short run. I formally estimate the discontinuity around the date the bill was signed using a bandwidth of +/- 24 months and various polynomials (linear, quadratic, cubic). The estimated discontinuity is never statistically different from zero, and it ranges from -5.5 to 2.1. Panel (b) plots the fraction of crack-cocaine seizures with 280-290g by race. The sample is limited to states with full coverage (i.e. all agencies in the state participating) starting in 2012 and with 90% coverage or more from at least 2008-2015.

Figure A11. Alternative Figures Testing for Shifting from State/Local Authorities to Federal Court

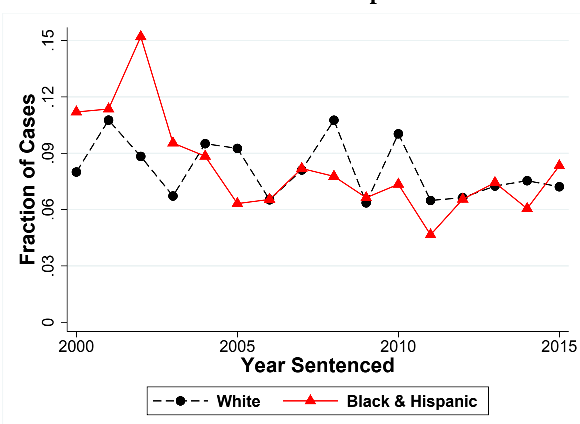
(a) Fraction of Cocaine Cases 200-400g, USSC



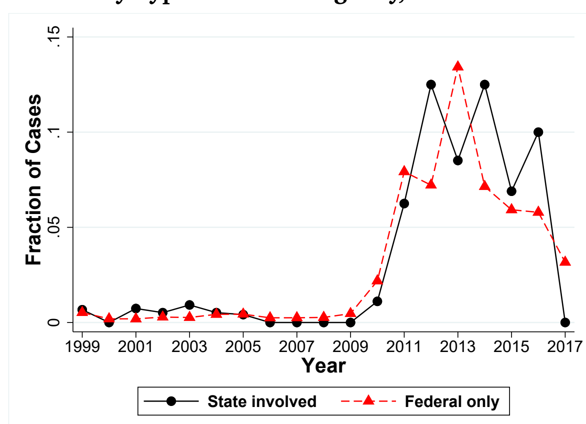
(b) Fraction of Cocaine Cases 200-400g, NC



(c) Fraction of Cocaine Cases 200-400g, FL, Alternative Sample

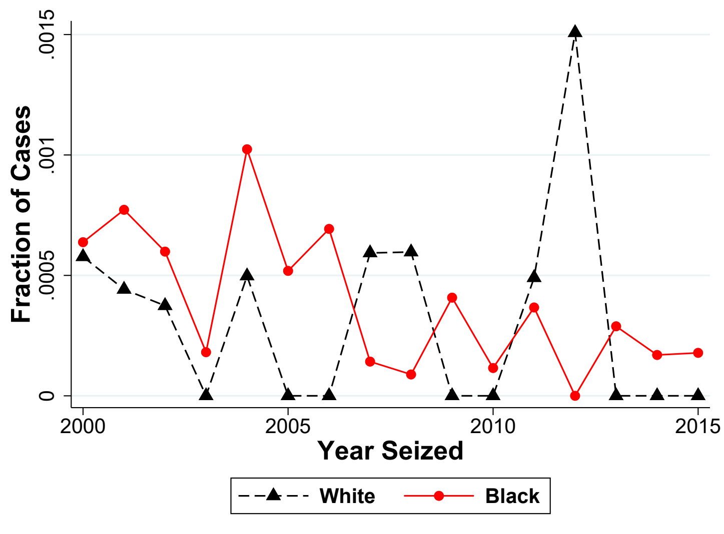


(d) Fraction of Crack-Cocaine Cases in 280-290g, by Type of Source Agency, EOUSA



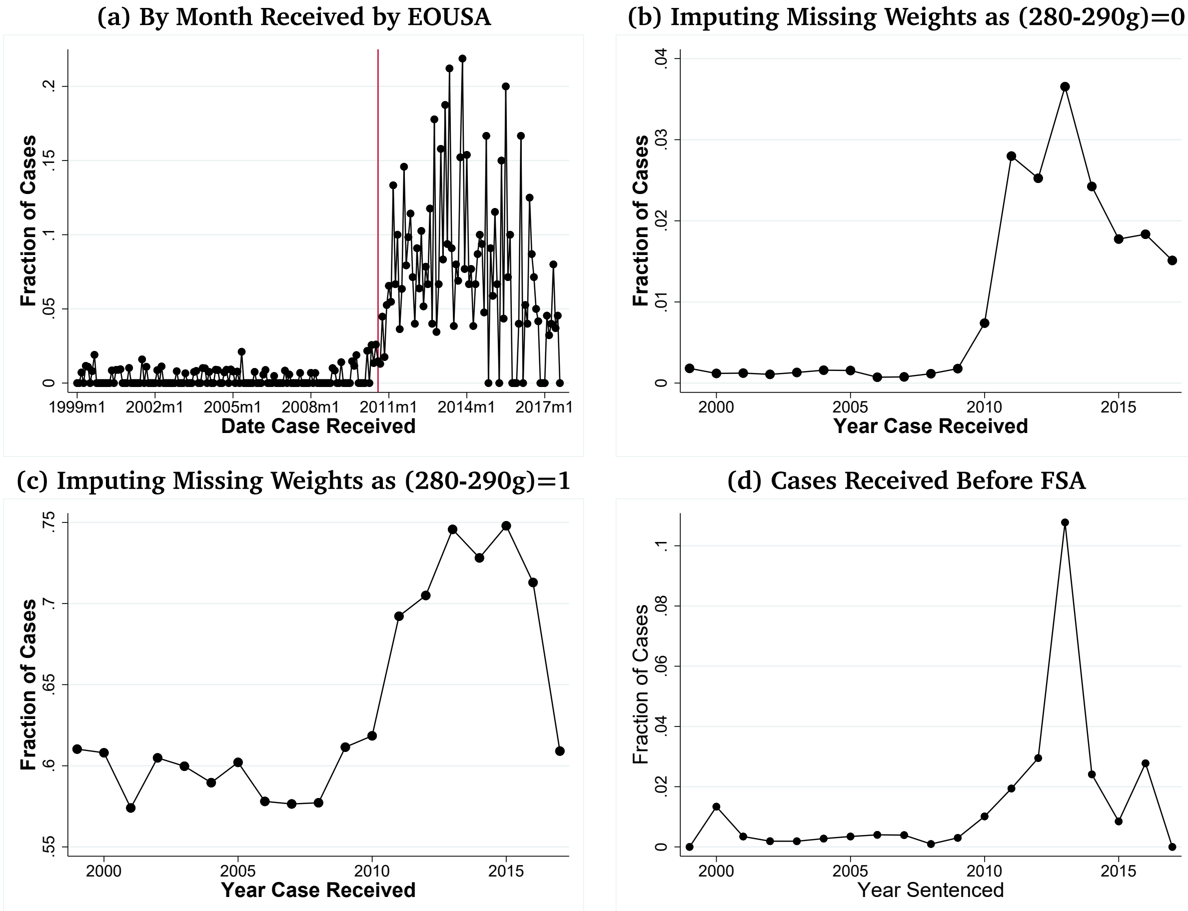
Notes. The figure in panel (a) plots the fraction of cocaine offenses with 200-400g in the USSC federal sentencing data, by race. The figure in panel (b) plots the fraction of cocaine offenses that have a range from 200-400g in NC state prison from 2000-2015, by race. Many of drug convictions in NC do not include type of drug in the offense description, the figure above is limited to those offenses that specifically list 'cocaine' in the offense description. The figure in panel (c) plots the fraction of cocaine offenses with 200-400g in FL state prison by race, limiting to those offenses that list a weight range in the offense description (the figure in the main text includes all cocaine offenses and codes (Convicted 200-400g)=0 if there is not weight listed in the offense description). The figure in panel (d) plots the share of cases sent to EOUSA attorneys from sources that involve state agencies (red dashed line with triangle markers) and the share of cases sent to EOUSA attorneys from strictly Federal sources (black solid line with circle markers). This figure is limited to the top agencies sending cases and excludes joint investigations (e.g. FBI + state/local task force). The top agencies are: DEA, FBI, ATF, and state/local.

Figure A12. Fraction of Crack-Cocaine Seizures from 280-290g, Full Coverage States, NIBRS



Notes. The figure above plots the fraction of crack-cocaine seizures with 280-290g by race. The sample is limited to states with full coverage (i.e. all agencies in the state participating) starting in 2012 and with 90% coverage or more from at least 2008-2015.

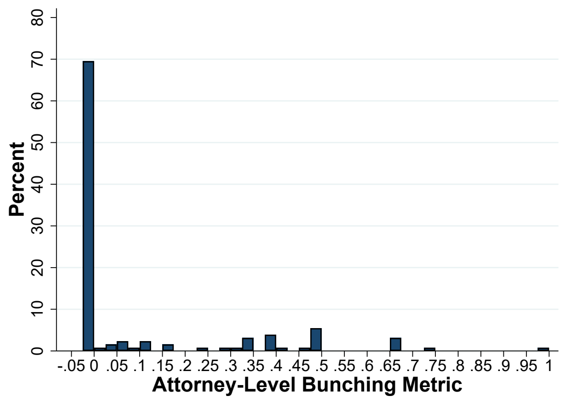
Figure A13. Fraction of Cases with 280-290g Over Time, EOUSA



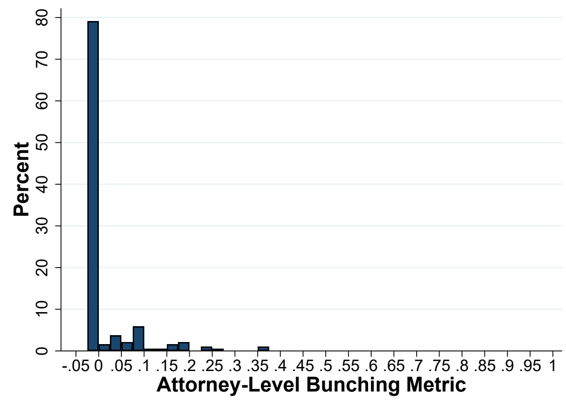
Notes. Panel (a) plots the fraction of cases with 280-290g (excluding cases with missing drug weights) by the month the case was received. The vertical red line indicates the date the Fair Sentencing Act was passed. In panel (b), I re-code the 280-290g dummy variable equal to zero if the drug weight is missing (typically, I leave the dummy variable missing if the drug weight is missing). In panel (c), I do the opposite, coding the 280-290g dummy variable equal to one if the drug weight is missing. In both cases, there is a sharp increase in the fraction of cases at 280-290g after 2010. Since panel (b) more accurately matches the statistics from the USSC final sentencing data, I use that imputed value for various robustness tests. Panel (d) plots the fraction of cases with 280-290g in each year for cases that are received by the EOUSA prior to the signing of the Fair Sentencing Act.

Figure A14. Histograms of Attorney-level Bunching Metric at 280-290g, EOUSA

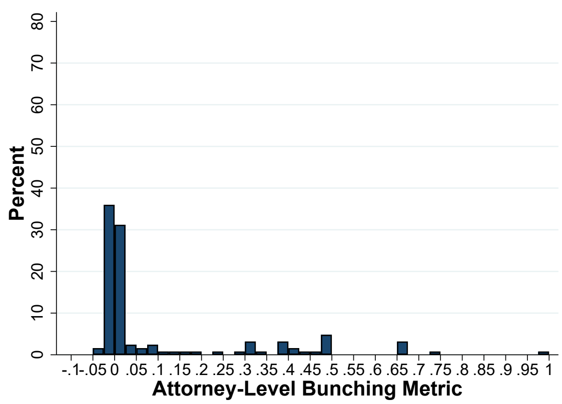
(a) 280-290g, National



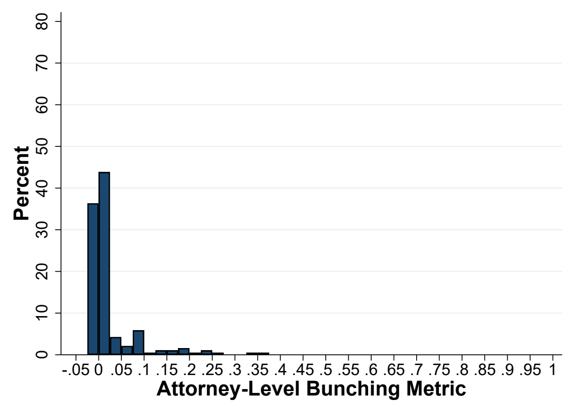
(b) 280-290g, Missing = 0, National



(c) 280-290g, District



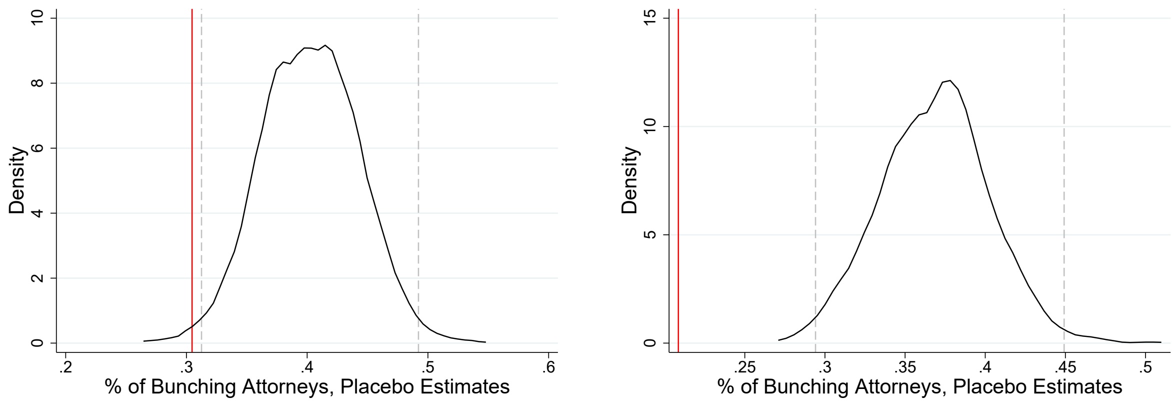
(d) 280-290g, Missing = 0, District



Notes. The figures above plot histograms of attorney-level bunching metrics, which are calculated as the difference between each attorney's fraction of cases with 280-290g post-2010 and the average fraction of cases with 280-290g at "baseline." In the national case (panels (a) and (b)), the baseline is the average fraction of cases with 280-290g prior to 2010. In the district case (panels (c) and (d)), the baseline for an attorney in district **A** is the average fraction of cases with 280-290g prior to 2010 in district **A**. Panels (b) and (d) include cases where the drug weight field is missing by coding the 280-290g dummy variable equal to zero when the drug weight is missing. I define an attorney as a "bunching attorney" if their bunching metric is above zero, thus the exact fraction of bunching attorneys for each panel is as follows: (a) 30.5%, (b) 20.9%, (c) 31.2%, and (d) 20.9%. These figures are limited to attorneys with 10+ cases post-2010. Limiting to 15+ cases delivers similar results. Limiting to 5+ cases decreases the fraction of bunching attorneys to: (a) 21.2%, (b) 14.2%, (c) 21.4%, and (d) 14.2%. Even imputing missing weight cases as though they **are** 280-290g cases (the highly unrealistic result in Figure A13c) implies that only 70% of attorneys bunch at 280-290g.

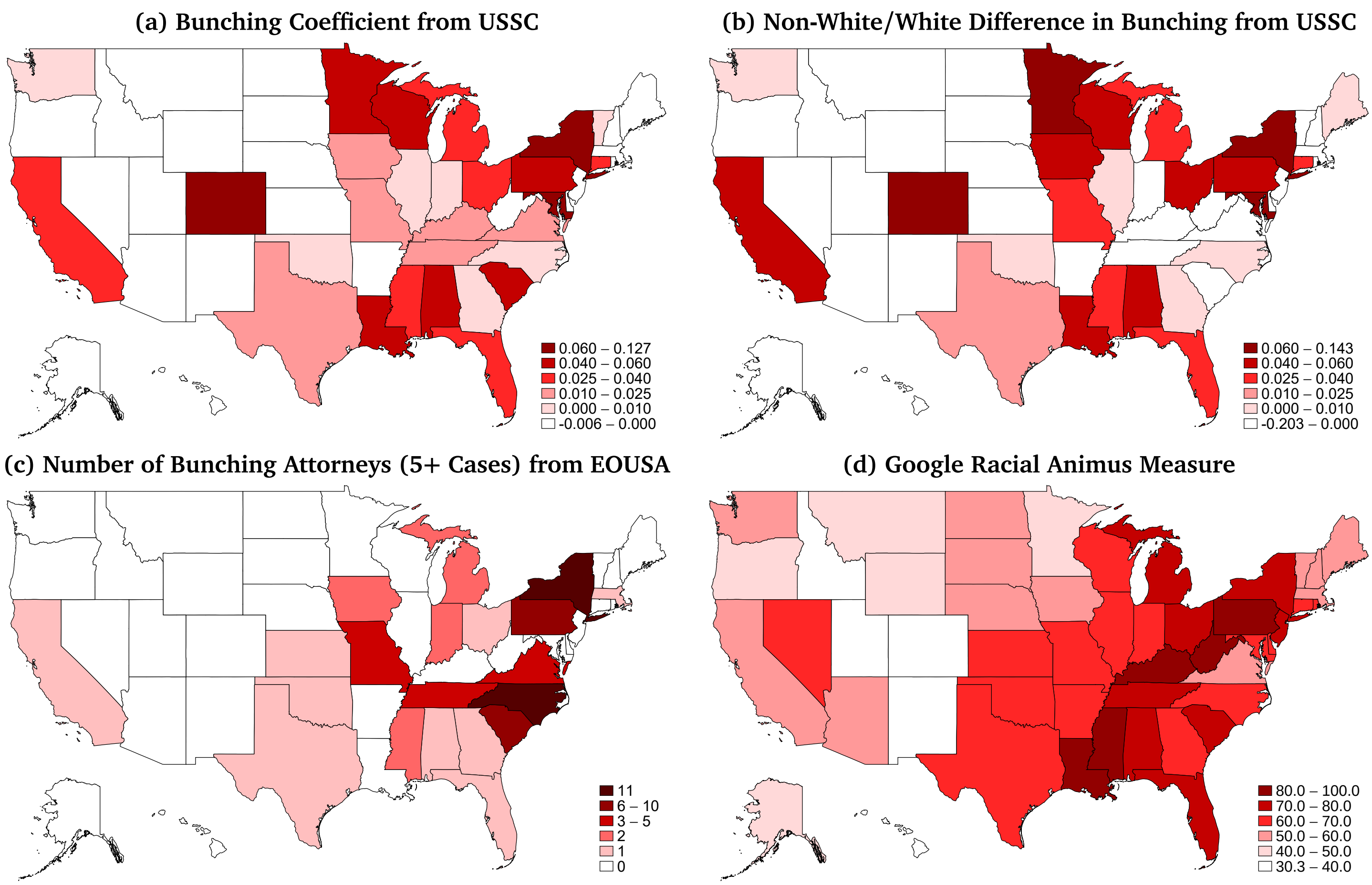
Figure A15. Histograms of Randomized Attorney-level Bunching Metric at 280-290g, EOUSA

(a) 280-290g, National **(b) 280-290g, Missing = 0, National**



Notes. I randomly re-assign all cases in the sample of attorneys with 10 or more cases after 2010, maintaining the same overall fraction of 280-290g cases in each year. After doing this random re-assignment, I calculate the number of bunching attorneys. I do this 1,000 times and plot the placebo estimates from the non-missing data in panel (a) and from the data with missing values imputed in panel (b). The gray dashed lines indicate the 1st and 99th percentiles of the placebo distribution and the red line indicates the fraction of bunching attorneys from the true data.

Figure A16. Map of State-level Bunching and State-level Racial Disparity in Bunching



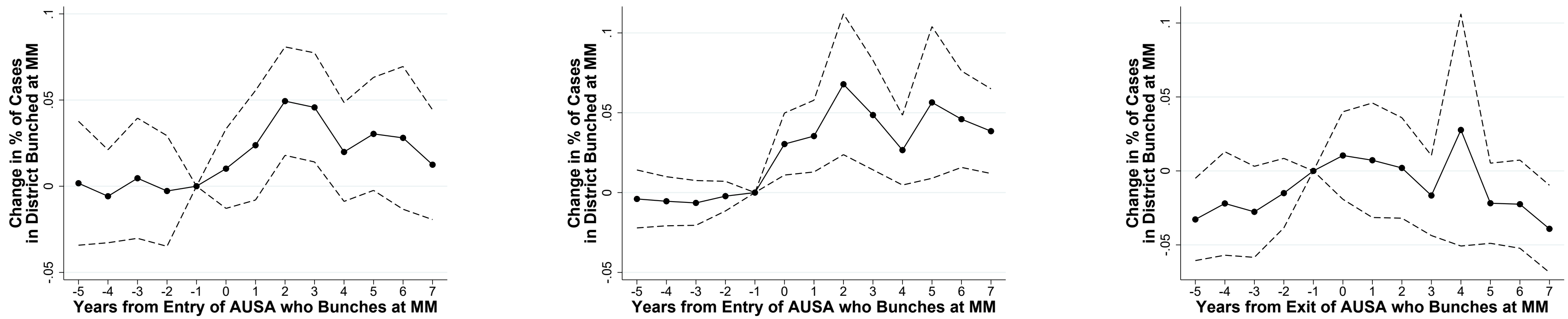
Notes. Panel (a) plots the state-level bunching estimate for all states with a sufficient number of cases. Panel (b) plots the difference between the state-level bunching estimate for white offenders and the state-level bunching estimate for black and Hispanic offenders for all states with a sufficient number of cases. Panel (c) plots the number of prosecutors who bunch in each state (among those prosecutors with 5+ drug cases after 2010). Panel (d) plots the racial animus index derived from Google search volume for a racial slur and introduced by Stephens-Davidowitz (2014). For Panels (a) and (b) there are several states that do not have enough cases to estimate bunching or racial disparities in bunching at 280-290g (these states are: AZ, DE, HI, ID, MT, ND, NH, NJ, NM, NV, OR, RI, SD, UT, WY). I pool all of these states in one regression and apply the resulting coefficient.

Figure A17. Additional Evidence of Prosecutorial Discretion in Bunching, Alleyne Results and Movers Results, EOUSA

(a) Effect of Entry of a Bunching AUSA

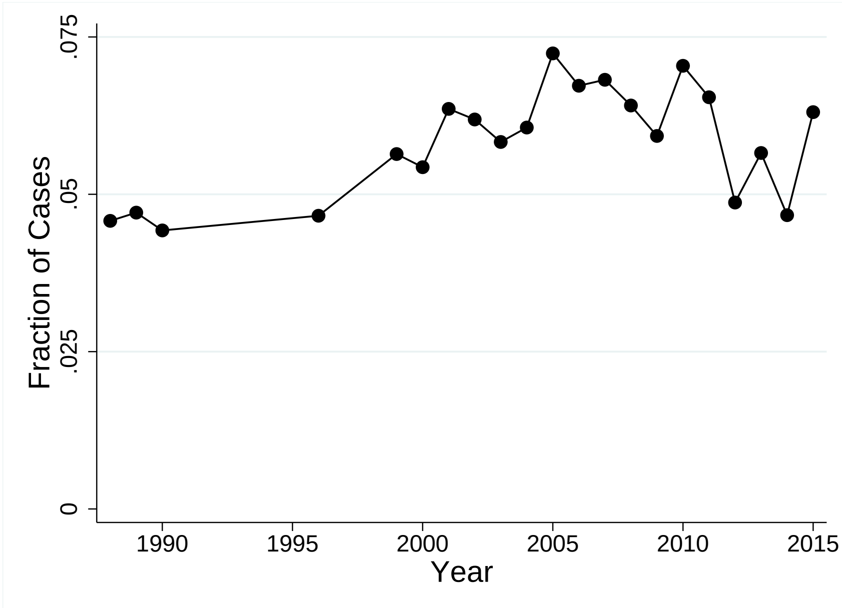
(b) Effect of Entry of a Bunching AUSA, Low-Bunching Districts

(c) Effect of Exit of a Bunching AUSA



Notes. Panels (a) and (b) plot the change in the percent of cases that are bunched at the mandatory minimum (MM) threshold (50g pre-2010 and 280g post-2010) after a "bunching" prosecutor enters a district. For these figures, I identify prosecutors who switch districts, who bunch at the mandatory minimum threshold in their first district, and who have 5 or more cases in their first district. I then identify the districts that they switch into and analyze the fraction of cases bunched at the mandatory minimum for all other prosecutors in that district. Panel (a) shows that prior to entry of a bunching prosecutor, district-level bunching does not change year-to-year, but that immediately after the bunching prosecutor enters, all other prosecutors in that district increase their fraction of cases bunched at the threshold. Panel (b) shows that this increase is driven by districts that have low-levels of bunching (below the median for all districts) prior to the entry of the bunching prosecutor. Panel (c) plots the bunching activity for the districts from which these prosecutors are leaving. This analysis is limited to the first bunching attorney from panels (a) and (b) that leaves the district. There is not a decrease in the prevalence of bunching after bunching prosecutors exit a district. This suggests bunching at the mandatory minimum threshold is not related to a temporary behavior shift, such as increased competition among attorneys, but that it may be related to something more permanent, such as learning about techniques or developing beliefs/norms. The dashed lines in panels (a)-(c) are 90% confidence intervals. Since these figures rely on prosecutors who move from one district to another and require reasonably long pre- and post-periods, I use data from 1994-2016 and identify the first moving attorney for post-1999 years only (insuring a 5-year pre-period for every district). In practice, this means the figures above are largely based on bunching at 50-60g (the pre-2010 mandatory minimum). Restricting to post-2010 moves does not yield a large enough sample of movers with sufficient cases to classify them as bunching versus non-bunching.

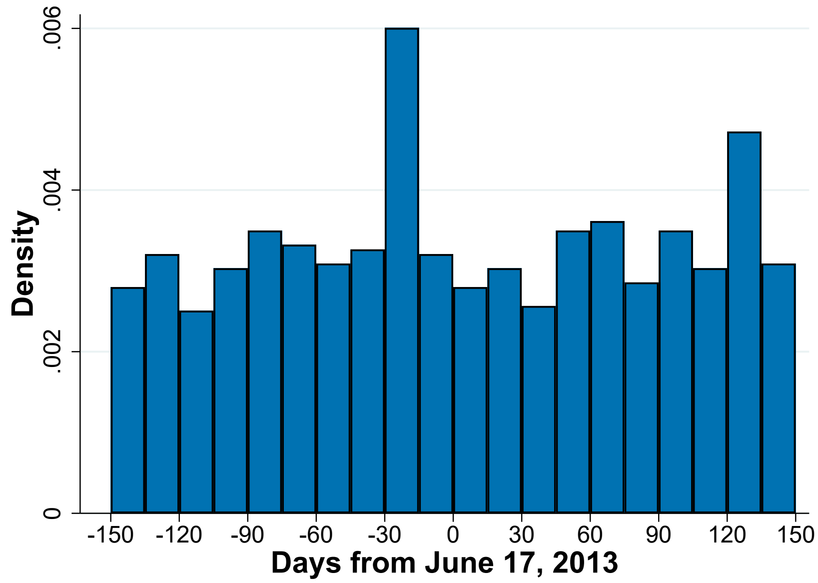
Figure A18. Fraction of Cases in 50-60g by Year, from USSC Sentencing Data



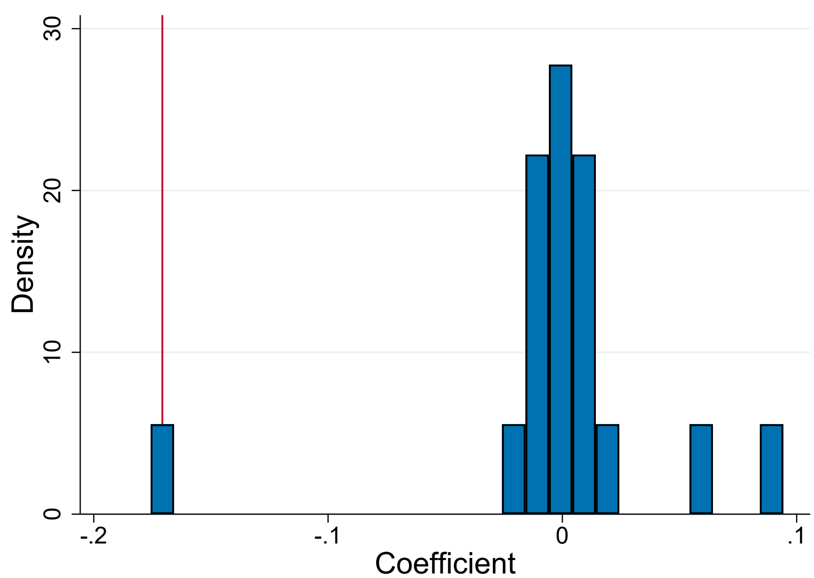
Notes. The figure above plots the fraction of all cocaine (powder and crack) cases with 50-60g by year. The sample is limited to cases with drug weights from 0-1000g. All cocaine cases are used because earlier years (1988-1990) do not distinguish between types of cocaine. This figure indicates that cases bunched above the pre-2010 10-year mandatory minimum threshold increased by about 60% from 1988-90 to 2010. Over this same time period, the average weight of cases from 0-1000g decreased. This suggests that the practice of bunching cases at the mandatory minimum was potentially learned over time, which is consistent with the evidence on movers and the spread of bunching in Figure A17.

Figure A19. Tests of Validity for Alleyne v. US Result, EOUSA

(a) Density of Cases Received Around June 17, 2013 (Date of Decision in Alleyne)

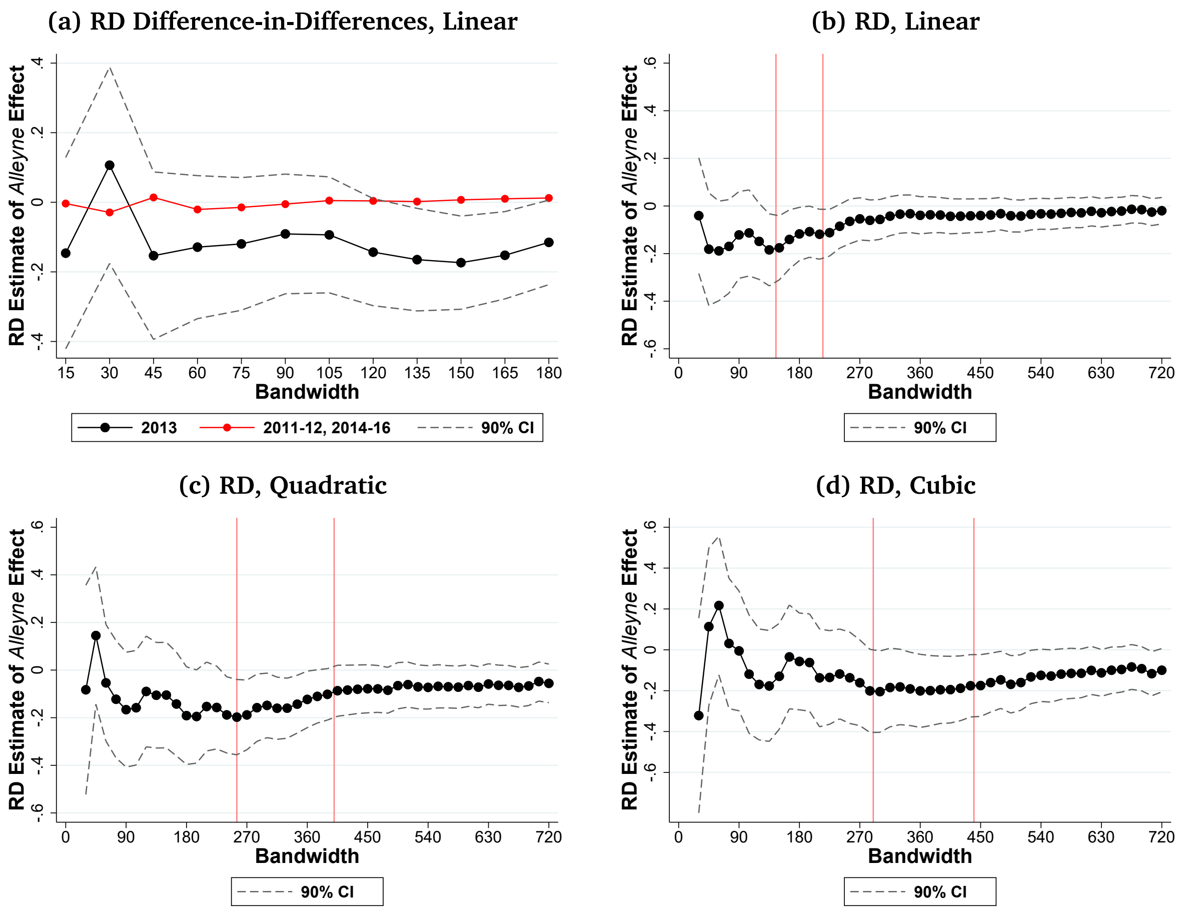


(b) Estimate of Discontinuity Around June 17 in All Years 1999-2016



Notes. Panel (a) plots the density of cases around the June 17, 2013 (centered at zero) and grouped into 15-day bins. June 17, 2013 is the day Alleyne v. US was decided. Outside of the large number of cases from -30 to -15 days before Alleyne was decided, the density is relatively smooth through that date. Panel (b) plots a histogram of the estimated discontinuity around June 17 in all years from 1999-2016. The estimates are centered at zero and the coefficient in June 2013 (marked by the red line) is twice as large as the next largest estimate of any sign and over 4 times larger than the next largest negative estimate.

Figure A20. Robustness of Alleyne v. US Result to Choice of Bandwidth and Polynomial, EOUSA



Notes. The figures above display estimates for the effect of Alleyne v. US (a case that strengthened evidentiary requirements) on the prevalence of bunching at 280-290g. Each panel displays estimates across many different bandwidth choices (i.e. the number of days before and after June 17 included in the regression) and different polynomial choices (i.e. the polynomial of the running variable, number of days from June 17, included in the regression) are shown across panels. Panel (a) displays coefficient estimates from the RD difference-in-differences regression for bandwidths from 15-180. Since the difference-in-difference estimates use multiple years, bandwidths above 160 days are asymmetric. The black line in panel (a) displays the estimates from 2013, the red line displays the estimates from all other years after 2010 (when nothing in particular happened around June 17). Panels (b)-(d) estimate a typical RD regression (i.e. not using variation around June 17 in other years). This allows me to extend the bandwidth to 2 years before and after Alleyne v. US. In these panels, the first red line denotes the CER-optimal bandwidth and the second red line denotes the MSE-optimal bandwidth (Cattaneo et al. 2018). In panel (b), for example, the estimate approaches zero at larger bandwidths–this is to be expected. As we get further from the cutoff, the a linear polynomial becomes an increasingly bad fit. In all three panels, the optimal bandwidths yield estimates that are statistically different from zero (or marginally statistically significant).

Appendix B. Alternative Methods of Estimating Bunching

I. Comparing Aggregated Pre- and Post-2010 Densities

Most papers using the "difference-in-bunching" approach can be fit into one of two categories. In one, authors estimate bunching using the conventional polynomial method (see section II below for a detailed description) separately for groups where the threshold applies and for groups where the threshold does not apply, using the latter as a placebo test (Best et al. 2015; Fack and Landais 2016; Gelber, Jones, and Sacks 2017; Zaresani 2017; Chen et al. 2018). In the other, authors directly compare the group where the threshold applies to the group where the threshold does not apply. Even within the direct comparison category, strategies differ. Several papers compare the distributions by aggregating the data into bins and calculating the difference in levels between the actual and the counterfactual distributions (Brown 2013; Best et al. 2018; Best and Kleven 2018; Cengiz, Dube, Lindner, and Zipperer 2018). Others compare the distributions using regression analysis on the microdata (Kleven et al. 2011; Behaghel and Blau 2012; Sallee and Slemrod 2012; Chetty, Friedman, and Saez 2013; Dwenger et al. 2016; Goncalves and Mello 2018; and Traxler et al. 2018). These papers frequently estimate the difference in the probability an observation is in a given bin between the actual and the counterfactual setting .

In this paper, I employ both direct comparison methods (aggregate/binned analysis and microdata analysis). I am primarily interested in estimating the change in the probability a case is charged with 280-290g after 2010 and whether that change in probability differs by race. In addition, some analyses in the paper preclude aggregating the data into bins because they rely on data that do not include precise drug quantities. For these reasons, I follow the papers that use regression analysis on microdata to compare the pre- and post-2010 crack-cocaine distributions.

To show robustness to the other "difference-in-bunching"/direct comparison method, I aggregate the cases into 10g bins pre- and post-2010. Following Best et al. (2018), I estimate 90% confidence intervals with a bootstrap procedure that samples cases with replacement from the microdata before aggregating to the 10g bin level.[1] I compare the binned distributions to estimate the net change in bins below 280g, at 280-290g, and above 290g.

Aggregate bunching analyses yield very similar results. Figure B1 below plots the counterfactual scaled pre-2010 density and the actual post-2010 density. The spike at 280g in the post-2010 density is the bunching that is detected in Table 2. After 2010, there is a 3.5 percentage point increase in cases with 280-290g. I also show the densities by race. The bunching at 280g in the post-2010 density is larger for black and Hispanic offenders. After 2010, the rise in cases with 280-290g is about 2 percentage points higher for black and Hispanic offenders than for white offenders.

In Figure B2a, I plot the difference between the post-2010 and the scaled pre-2010 densities for each 10g bin and add confidence intervals by using 50 bootstrapped samples from the microdata. In addition, I also display a table of the statistical results for the binned missing mass analysis in Figure B2b. When this difference is below zero, it means the bin contains relatively fewer cases after 2010 and when the difference is above zero, it means the bin contains more cases after 2010.

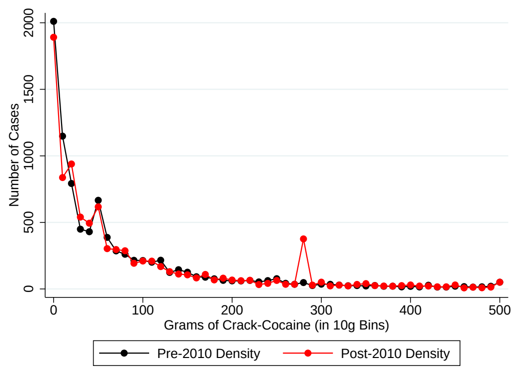
The figure shows an increase of about 340 cases in the 280-290g bin post-2010, a net increase in cases above 280g, and a net decrease below 280g. Summing the changes in bins above 280g, I find a net increase in that section of the distribution after 2010. The point estimate on the net change is noisy, but even summing the lower bound of the confidence interval for all bins above 280g can only account for about 46% of the increase in the 280-290g bin. On the other hand, the net change below 280g can account for 120% of the increase in the 280-290g bin. Again, this point estimate is noisy. In fact, summing the upper confidence interval for all bins below 280g implies

[1] I draw 50 random samples from the microdata and do the binned analysis on each sample. The final number of cases for each bin is calculated as the mean of the number of cases across all 50 samples, and the final standard error is calculated as the mean of the standard error across all 50 samples.

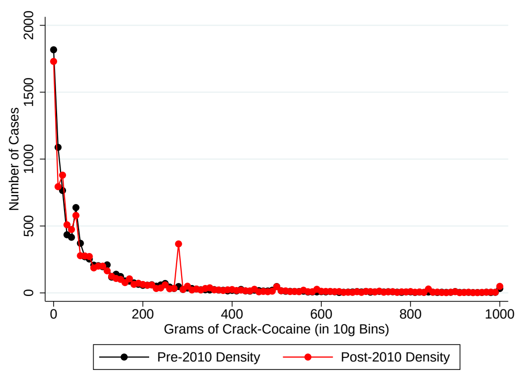
a net increase in that section of the distribution. The key takeaway is that changes in the distribution below 280g can account for the excess mass at 280g, whereas changes in the distribution above 280g cannot. In other words, an offender charged with 280-290g post-2010 would likely have been charged with less than 280g had they been sentenced prior to 2010. Table B1 displays the results from similar binned analyses using the NIBRS data, DEA data, and EOUSA data.

Figure B1. Scaled Pre-2010 Distribution of Recorded Weights vs. Post-2010 Distribution

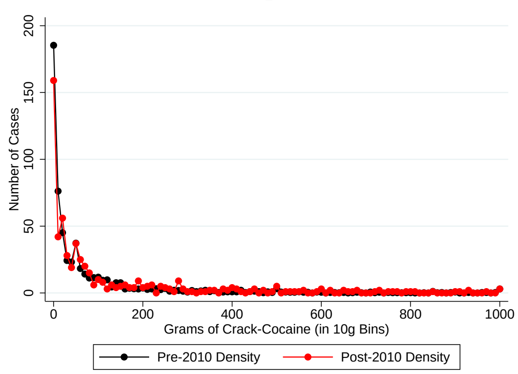
(a) All Offenders



(b) White Offenders



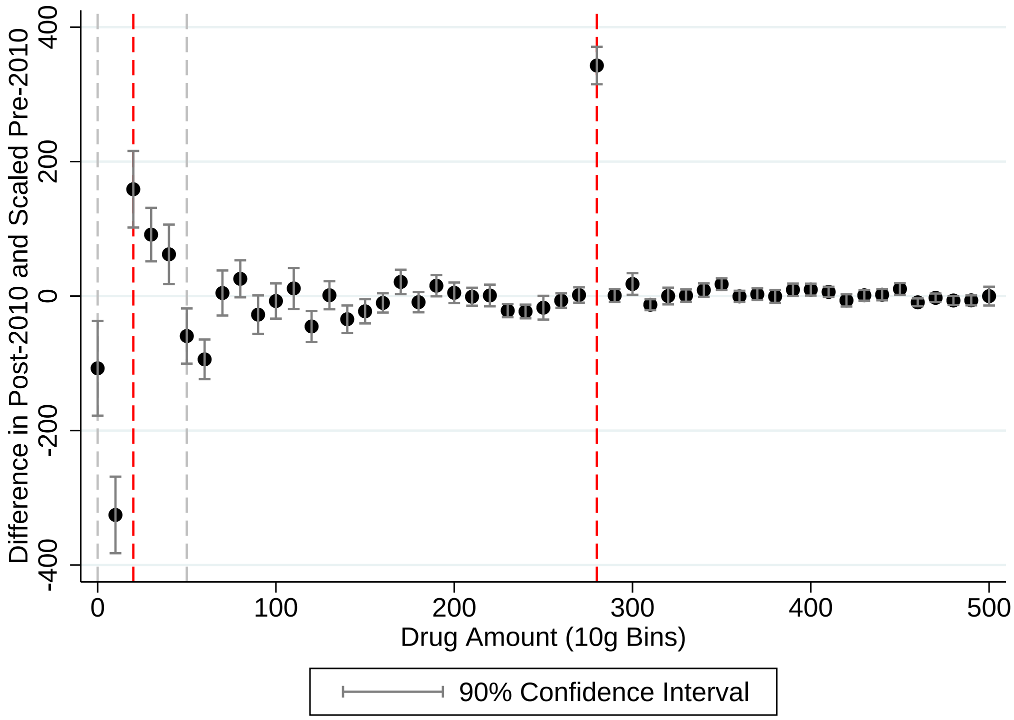
(c) Black and Hispanic Offenders



Notes. Figure B1a plots the scaled density of drug quantities pre-2010 (in black) and the actual density of drug quantities post-2010 (in red) for all offenders. The amounts are aggregated into 10-gram bins and limited to drug quantities under 1000g. Figures B1b and B1c do the same but restrict the sample to white offenders or black and Hispanic offenders, respectively.

Figure B2. Post-2010 Density Minus Scaled Pre-2010 Density

(a) Difference between Post-2010 and Pre-2010 Distribution of Drug Amounts



(b) Fraction of Bunching Accounted for by Different Ranges

Range	Net Difference	90% CI	% Bunching at 280g
0-20g	-435.56	(-558.17, -312.94)	128.67%
20-50g	293.49	(158.10, 428.88)	-86.70%
50-60g	-52.63	(-101.67, -3.59)	15.55%
60-100g	-65.00	(-184.58, 54.59)	19.20%
100-280g	-122.43	(-414.25, 169.39)	36.17%
0-280g	-382.13	(-1100.58, 336.32)	112.89%
290-500g	43.44	(-146.74, 233.62)	-12.83%

Notes. The figure above plots the difference between the post-2010 density and the scaled density of drug quantities in pre-2010 for each 10-gram bin. Confidence intervals are calculated by bootstrapping as discussed in the text. The red dashed lines correspond to the post-2010 mandatory minimum bins (28g and 280g) and the gray dashed lines correspond to the pre-2010 mandatory minimum bins (5g and 50g). Summing the changes in bins above 280g, I find a net increase in that section of the distribution after 2010. The point estimate on the net change is noisy, but even summing the lower bound of the confidence interval for all bins above 280g can only account for about 46% of the increase in the 280-290g bin. On the other hand, the net change below 280g can account for 120% of the increase in the 280-290g bin. Even the changes from 50-280g can account for 85% of the increase in the 280-290g bin. Panel B displays statistical results for relevant drug amount ranges.

Table B1. All Bunching Results using Aggregated/Binned Comparison with Bootstrapped SEs

	Pr(280-290g Crack-Cocaine Recorded)					
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0347***		-0.0002***		-0.0006***	0.0771***
	(0.0020)		(0.0001)		(0.0002)	(0.0054)
After 2010 x White		0.0126**		-0.00002		
		(0.0062)		(0.0001)		
After 2010 x Non-White		0.0359***		-0.0003***		
		(0.0023)		(0.0001)		
Constant	-0.0003***	-0.0001***	0.000002	0.0000002***	0.000006***	-0.0008***
	(0.00002)	(0.0001)	(0.0000008)	(0.000001)	(0.000002)	(0.0001)
Data	USSC, Final Sentencing	USSC, Final Sentencing	NIBRS, Drug Seizures	NIBRS, Drug Seizures	DEA, Drug Seizures	EOUSA, Prosecutor Files
Bins	100	100	100	100	100	100
Observations	57,101	52,940	203,700	203,700	100,306	24,493

Notes. Bootstrapped standard errors in parentheses. Standard errors are calculated from the standard deviation in estimates derived from 50 replications where in each replication cases are sampled with replacement before aggregating to the 10g bin level. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. Specifications with the white/non-white and after 2010 interactions also include a dummy variable equal to one for black and Hispanic offenders. Columns 1-2 show the main bunching result for the final sentencing data. Columns 3-5 show no increase in bunching for drug seizure amounts. Column 6 shows an increase in bunching in prosecutor case management files.
*** p<0.01, ** p<0.05, * p<0.1

II. Comparing an Estimated Counterfactual and Post-2010 Densities

Many bunching papers, for lack of variation in the threshold of interest, estimate bunching by constructing the counterfactual density from the actual bunched density. To do this, one typically aggregates the data into bins and estimates a regression of the count in each bin on a high-order polynomial of the bin's value and dummy variables for bins in the bunched "window." The estimates from that regression (not including the bunching dummy variables) can be used to predict a smooth distribution of bin counts. Authors then compare that smooth density to the actual density to calculate the degree of bunching in the actual density. My main results are also robust to this method.

To start, I collapse the data on drug quantities for all cases after 2010 to 10 gram bins. I then run a regression of the count of cases on a seventh order polynomial of the bin values and dummy variables for the bins 0-10g, 270-280g, and 280-290g. Then, using the coefficients from the seventh order polynomial and the dummy variable for the bin 0-10g, I calculate a smooth counterfactual distribution. For graphical purposes, I re-scale that smooth distribution to have the same total number of cases as the true distribution. Next, I calculate the percent of all cases that are in the 280-290g bin in the true distribution, the percent of all cases that are in the 280-290g bin in the counterfactual distribution, and the difference between those two percentages. Finally, I run a regression of the difference between the true and counterfactual distributions on a dummy variable equal to one for the 280-290g bin and equal to zero otherwise (bootstrapped standard errors are calculated by re-sampling the residuals from the polynomial estimation with 200 replications). I carry out a similar procedure to estimate the difference in bunching between white and black and Hispanic offenders (the major difference being that I estimate the counterfactual distributions separately for white and black and Hispanic offenders and that the final regression includes an interaction between the 280-290g bin dummy and a dummy for black and Hispanic offenders).

First, I construct the counterfactual density by aggregating the data to 10-gram bins, summing the number of cases in each bin. With this aggregated data, I estimate a regression of the bin counts on a seventh-order polynomial of the bin values, dummies for the 270g and 280g bins, and a dummy for the 0g bin.

$$Count_b = \alpha_0 + \sum_{i=1}^{7} \beta_i (Amount_b)^i + \gamma_1 Bin270_b + \gamma_2 Bin280_b + \delta_1 Bin0_b + \epsilon_b \tag{1}$$

where $Count_b$ is the total number of cases in bin b, $Amount_b$ is the value of bin b, and $Bin[X]_b$ is a dummy variable indicating if the bin's value equals X. I use the parameter estimates from (8) (excluding γ_1 and γ_2) to predict a smooth density of bin counts. Furthermore, I adjust the predicted counts to force the smooth density to have the same number of cases as the actual density. I plot the counterfactual density and the actual post-2010 density below in Figures B3 and B4.

Using the predicted counts from the counterfactual density and the actual counts post-2010, I construct the percent of cases in each bin for each density. I then calculate the difference in these percentages and run the following regression, bootstrapping the standard errors from 200 replications:

$$(\text{\% in Post2010 - \% in Predicted})_b = \alpha + \beta Bin280_b + \epsilon_b$$

The resulting $\beta = 0.0352$ and $SE_\beta = 0.0169$.

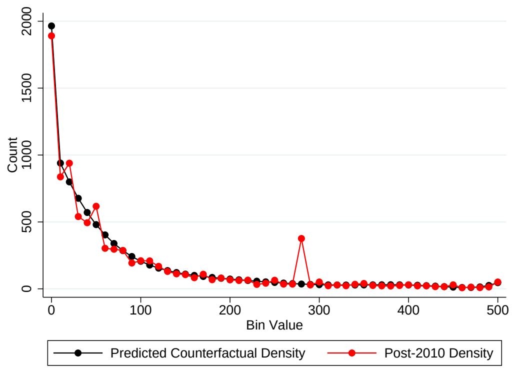
Next, I estimate:

$$(\text{\% in Post2010 - \% in Counterfactual})_{br} = \alpha + \beta Bin280_b + \gamma NonWhite_r + \delta Bin280_b \times NonWhite_r + \epsilon_b$$

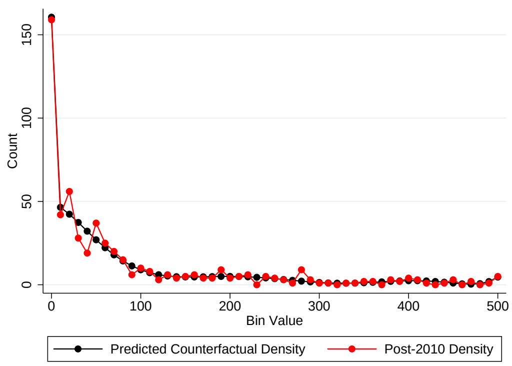
Using the Saez (2010) and Chetty et al. (2011) method, I estimate $\delta = 0.0237$ and $SE_\delta = 0.0119$. Using the difference-in-bunching method, I estimate $\delta = 0.0216$ and $SE_{\delta} = 0.0109$. In all analyses, I detect substantial bunching after 2010 and disproportionate bunching after 2010 for black and Hispanic offenders.

Figure B3. Predicted Counterfactual Density and Post-2010 Density

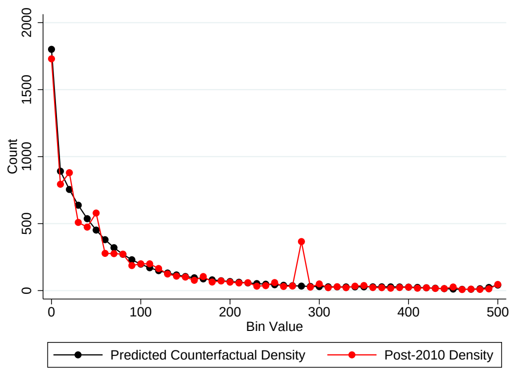
(a) All Offenders



(b) White Offenders, Saez (2010) Method



(c) Black and Hispanic Offenders, Saez (2010) Method



Notes. In panel (a), I plot a predicted counterfactual density of drug quantities (in black) and the actual density of drug quantities post-2010 (in red). In panels (b) and (c), I plot predicted counterfactual densities of drug quantities (in black) and the actual densities of drug quantities post-2010 (in red) by race. The amounts are aggregated into 10-gram bins and limited to drug quantities under 500g.

References

Behaghel, Luc, and David M Blau. 2012. "Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age." *American Economic Journal: Economic Policy* 4 (4): 41–67.

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy* 123 (6): 1311–55.

Best, Michael and Henrik Jacobsen Kleven. 2018. "Housing Market Responses to Transaction Taxes: Evidence from Notches and Stimulus in the UK." Working paper.

Best, Michael, James Cloyne, Ethan Ilzetzki, and Henrik Jacobsen Kleven. 2018. "Interest Rates, Debt and Intertemporal Allocation: Evidence from Notched Mortgage Contracts in the United Kingdom." Working paper.

Brown, Kristine M. 2013. "The Link between Pensions and Retirement Timing: Lessons from California Teachers." *Journal of Public Economics* 98 (February): 1–14.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2018. "The Effect of Minimum Wages on Low-Wage Jobs: Evidence from the United States Using a Bunching Estimator." Working paper.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suarez Serrato, and Daniel Yi Xu. 2018. "Notching R&D Investment with Corporate Income Tax Cuts in China." Working paper.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics* 126 (2): 749–804.

Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103 (7): 2683–2721.

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 2016. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." *American Economic Journal: Economic Policy* 8 (3): 203–32.

Fack, Gabrielle, and Camille Landais. 2016. "The Effect of Tax Enforcement on Tax Elasticities: Evidence from Charitable Contributions in France." *Journal of Public Economics* 133 (January): 23–40.

Gelber, Alexander M., Damon Jones, and Daniel W. Sacks. 2013. "Estimating Earnings Adjustment Frictions: Method and Evidence from the Social Security Earnings Test." NBER working paper.

Goncalves, Felipe and Steven Mello. 2018. "A Few Bad Apples? Racial Bias in Policing." Working paper.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Soren Pedersen, and Emmanuel Saez. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." 2011. *Econometrica* 79 (3): 651–692.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics* 8: 435–464.

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3): 180–212.

Sallee, James M., and Joel Slemrod. 2012. "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." *Journal of Public Economics* 96 (11): 981-999.

Traxler, Christian, Franz G. Westermaier, and Ansgar Wohlschlegel. 2018. "Bunching on the Autobahn? Speeding Responses to a 'Notched' Penalty Scheme." *Journal of Public Economics* 157: 78–94.

Zaresani, Arezou. 2016. "Adjustment Costs and Incentives to Work: Evidence from a Disability Insurance Program." University of Calgary working paper.

Appendix C. Supplementary Materials for Model of Prosecutor Objectives

I. Prosecutor Responses to Changing Mandatory Minimum Thresholds

The model from Section V.E also has implications about how the optimal choice in period $t = 0$ relates to the optimal choice in period $t = 1$. I outline this in Section II.B.2, and provide additional detail in this Appendix section.

Assuming that there are no fixed costs to building a case and that there are no changes in the objective function other than the change in the sentencing schedule, then a prosecutor who chooses not to bunch a case at a mandatory minimum threshold for a sentence X in one period would not bunch the same case at a higher mandatory minimum threshold for a sentence $Y \leq X$ in another period. In other words, a prosecutor not taking on the costs of bunching for a given gain would not take on even greater costs for the same or lesser gain.

For example, when a prosecutor chooses $a^{0*} = s < 5$, this implies that their utility from choosing s is higher than their utility of choosing 5g or 50g: $u(s)^0 > u(5)^0$ and $u(s)^0 > u(50)^0$. Since $a^{1*} = 28$ yields the same benefits as $a^{0*} = 5$ but requires greater costs, then $u(5)^0 > u(28)^0$. These two statements (and the assumptions above) imply that $u(s)^1 > u(28)^1$, which means that the prosecutor should also choose $a^{1*} = s < 5$. The same revealed preference argument can be made for why $u(s)^1 > u(280)^1$. Table 1 shows these possible rational choices of a^{1*} for a given a^{0*} and ranges of s.

Table 1. Relationship between a^{0*} and a^{1*} for relevant ranges of seized evidence

	(1) $s < 5$	(2) $28 > s \geq 5$	(3) $50 > s \geq 28$	(4) $280 > s \geq 50$	(5) $s \geq 280$
$a^{0*} = s$	$a^{1*} = s$	$a^{1*} = \{s, 28\}$	$a^{1*} = s$	$a^{1*} = \{s, 280\}$	$a^{1*} = s$
$a^{0*} = 5$	$a^{1*} = \{s, 28\}$	–	–	–	–
$a^{0*} = 50$	$a^{1*} = \{s, 28, 280\}$	$a^{1*} = \{s, 28, 280\}$	$a^{1*} = \{s, 280\}$	–	–

Ultimately, this means that there will be an increase in the share of cases with $a < 5$ post-2010 (increases from cases previously bunched at 5g and 50g); an ambiguous change in the share of cases with $28 > a \geq 5$ (increases from cases previously bunched at 50g and decreases from cases previously bunched at 5g), an increase in the share of cases with $50 > a \geq 28$ (increases from cases previously bunched at 5g and cases previously bunched at 50g); a decrease in the share of cases with $280 > a \geq 50$ (decreases from cases previously bunched at 50g and cases previously left with $a = s \geq 50$), and an increase in the share of cases with $a \geq 280$ (increases from cases previously bunched at 50g and cases previously left with $a = s \geq 50$). See Figure A2 for a graphical representation of this.

II. Prosecutors' Signal Extraction Problem

The racial disparity in bunching at 280g after 2010 could be due to statistical discrimination. Recall that seized evidence s is a noisy measure of true drug trafficking d. Suppose that, on average, black and Hispanic defendants have higher true drug trafficking amounts:

$$
\begin{aligned}
d_r &\sim N(\bar{d}_r, \sigma_d^2) \\
\bar{d}_{bh} &> \bar{d}_w
\end{aligned}
$$

Since s is a noisy measure of true drug trafficking d, we can write s as follows:

$$s = d + \nu,\ \nu \sim N(\mu, \sigma_\nu^2)$$

This implies that $E(d|s,r,x) = \bar{d}_r \times (1-\alpha) + (s-\mu) \times \alpha$ where $\alpha = \sigma_\nu^2/(\sigma_\nu^2+\sigma_d^2)$. Since $\bar{d}_{bh} > \bar{d}_w$, $E(d|s,bh,x) > E(d|s,w,x)$. Since the prosecutor does not observe d, they instead use $l^*(E(d|s,r,x))$. I denote this as $l^*(s,r,x)$, and the setting described here implies that $l^*(s,bh,x) > l^*(s,w,x)$. In other words, the prosecutor's expectation over true drug trafficking d "justifies" a higher sentence for black and Hispanic offenders. This decreases their cost of choosing $a > s$ because the associated mandatory minimum sentence will be less of a deviation from that sentence l*. Prosecutors may also use another defendant characteristic x_1 to solve the signal extraction problem (as detailed above) and arrive at $l^*(s,r,x=x_1) > l^*(s,r,x \neq x_1)$.

Appendix D. Data Appendix

United States Sentencing Commission (USSC) Federal Sentencing Data

These data contain the universe of federal sentences from 1999-2015. The data were obtained from the ICPSR "Monitoring of Federal Criminal Sentences" series here:
https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/83.
The data itself is compiled from several court documents: (1) the Judgment and Conviction Order (JC), (2) the Pre-sentence Report (PSR), and (3) the Statement of Reasons (SOR). The PSR is prepared by the probation officer in consultation with the prosecutor and the defense. It is a detailed report on the offender and their offenses intended to aid the judge in making the factual determinations that affect sentencing. The SOR is a form filled out by the judge that details their findings and whether/why they differ from the PSR. The JC is the final ruling in the case that outlines the adjudication and the sentence. Key variables from the data are described below:

Crack cocaine offense. Whether or not the case involves a crack cocaine offense is derived from the raw variables DRUGTYP{X} provided by USSC. These variables contain the types of drugs involved in the offense. This information is taken from the Judgment and Conviction Order (JC), if present. If it is not included in the JC, the information is taken from the Pre-sentencing Report (PSR) prepared by the probation officer assigned to the case. According to USSC, if the information in these documents conflicts, the JC takes precedent.

Drug quantity. The amount of drugs involved in the case is derived from the variables WGT{X} provided by the USSC. These variables contain the gram amount for drug {X} corresponding to DRUGTYP{X}. I use the weight corresponding to the drug type crack cocaine for each case. The values for WGT{X} are converted from variables DRGAM{X} and UNIT{X}. Information on drug amount and drug unit is taken from the Statement of Reasons (SOR), if present. If not present in the SOR, the information is taken from the PSR. According to the USSC, if the information in these documents conflicts, the SOR takes precedent.

Offender race. I code offender race based on the USSC variables NEWRACE, which categorizes offenders as non-Hispanic white, non-Hispanic black, or Hispanic. The variable NEWRACE is a combination of raw variables MONRACE and HISPORIG. The information for these variables is taken from the PSR. In fact, the USSC notes that offender race is self-reported to the probation officer.

Other offender characteristics (e.g. education). These are also derived primarily from the PSR.

Year. The year used for analyses is derived from the variable AMENDYR, which represents the year of the guideline manual used for sentencing guidelines calculations. This information is taken from the PSR.

District. The district used for analyses is derived from the variables DISTRICT, which represents the federal district the offender is sentenced in. This information is taken from the JC, if available, and from the PSR, if not. If both documents are available, and the information conflicts, the JC takes precedent.

FL State Inmate Database

These data contain all inmates who have been released from a FL state prison since October 1997. The data were obtained here: http://www.dc.state.fl.us/pub/obis_request.html. Key variables from the data are outlined below:

Offense/drug quantity. The offense field indicates all of the inmate's known offenses in FL. For drug offenses, the field contains the drug name. In FL, powder-cocaine and crack-cocaine cases are both recorded as "cocaine." For many of the drug offenses, the field contains a label indicating if the offense was with 0-28g of cocaine, 28-200g, 200-400g, or 400+g.

Offender race. Offender race is included as part of the "basic inmate information" file. There is no information on how race is determined. I expect it is similar to the federal court data, in which race is self-reported. In the FL data, the race field includes labels for "black", "Hispanic", and "white" inmates.

In robustness tests, I use similar data from North Carolina. It also contains an offense string that provides information about drug type and quantity. However, the string does not always specific the type of drug. These data cover cases that are handled at the state/local level as opposed to federal court (those cases included in the USSC data). This is important because state and local authorities could send more of their high weight, 280g cases to federal court after 2010. Similarly, federal prosecutors could pull more of these types of cases from state and local courts after 2010. A case can enter the federal system for procedural reasons: drugs are trafficked across state lines or the arrest is made by federal agents. However, cases can also be prosecuted federally for more arbitrary reasons. Wright (2006) notes that sorting into federal versus state is often determined by law enforcement agents involved with the case and/or the prosecuting attorneys, but it is never the official purview of judges or defense attorneys.[1] Why might local law enforcement or attorneys wish to pass a case on to the federal courts? For one, local authorities may not have the time or resources to properly pursue a case. Also, Wright suggests that federal sentencing is typically harsher than state sentencing, and that this gap could motivate jurisdiction decisions.

NIBRS Property Segment

These data contain information on drug quantity and drug type for drugs seized by NIBRS-participating police departments. The data were obtained here: icpsr.umich.edu/icpsrweb/NACJD/series/128. Key variables from the data are outlined below:

Drug quantity. The drug quantity field is populated when there is a drug seizure by the department. It is equal to the total quantity of drugs seized.

[1] **Wright, Ronald.** 2006. "Federal or State? Sorting as a Sentencing Choice." *Criminal Justice* 21 (2): 16-21.

Offender race. The race field for NIBRS does not include an indicator for whether the offender is Hispanic. An ethnicity field is available only in later years, so I focus on white versus black offenders in this data. There is no information on how race of the offender is determined. I expect it is similar to other criminal justice data, in which race is self-reported.

For the primary analyses of the NIBRS data, I limit the sample to a balanced panel of agencies. For robustness checks, I limit to stats that have had full agency coverage in NIBRS since 2012 and over 90% coverage since 1998.

DEA STRIDE Database

These data contain information on drug quantity, drug type, and purity for seizures and undercover purchases sent to DEA labs for analysis. The data also indicate whether the drugs were obtained via seizure or undercover purchase. For drugs that were purchased, the data contains their price. The data were obtained from a FOIA request for all records related to cocaine from January 1, 1999 to December 31, 2015. Key variables from the data are:

Drug quantity. This field indicates the weight of the drug evidence received by the lab.

Drug type. This field indicates type of drug. The DEA does not use street names to refer to drugs in this data, meaning no drugs are referred to as crack-cocaine. For the main analyses, I use all drug types containing the word "cocaine," but results are similar if I focus on the "cocaine base" drug type.

Purity. This field indicates the chemical purity of the drug evidence received by the lab.

Acquisition. This field indicates whether the drug was acquired via seizure or undercover purchase.

Price. This field is populated if the drugs were acquired via undercover purchase. Price indicates the price paid for the drugs. In one robustness analysis, I plot the time series of price by month. To do this, I adjust the raw price field (described here) based on the purity of the drug, calculating a "price per pure gram."

EOUSA Case Management Files

These data contain information on cases handled by the EOUSA from the EOUSA's internal case management system: Legal Information Office Network System (LIONS). The data were obtained here:
https://www.justice.gov/usao/resources/foia-library/national-caseload-data.
Key variables from the data are:

Drug quantity/type. This field comes from the "controlled substances" screen of the LIONS software. According to the LIONS user manual, the controlled substances data "tracks information on controlled substances; includes type and quantity of all substances in a case." The manual instructs users to do the following: "Enter the actual quantity of the controlled substance seized. Fractions must be converted to one or two decimal places." The software itself, however, simply has a field for "quantity" to be entered with no instruction. In general, the drug weights recorded in the EOUSA data are much larger than the drug seizure weights reported by the DEA or NIBRS. In fact, drug quantities decrease in the DEA and NIBRS after 2010 but increase in the EOUSA. Also, the fraction of 280-290g cases at the district/month level in the EOUSA data is highly correlated with the fraction of 280-290g cases at the district/month level in the USSC data. These validation tests suggest the data entered into LIONS is indicative of total drugs involved/charged in the offense and not raw amount seized alone.

Staff ID/Assignment. The EOUSA data also contains an ID variable for the lead attorney assigned to the case. This ID is tied to the district. In other words, two attorneys can have the same numeric ID as long as they are in different districts. Also, this ID will not follow an attorney from one district to another.

Initials. Since the EOUSA numeric ID for lead attorney is not constant across districts, I use a field for the attorney's "initials" to follow attorneys who switch districts. The initials field is "initials of the staff member authorized to use the LIONS application." In most cases, the field contains 3 or more letters, making it likely that if I see the same initial in two different districts it is the same attorney. In practice, this initials-based ID appears to accurately identify attorneys who switch districts. First, attorneys who move from one district to another continue to bunch at 280g in the new district. Second, when an attorney moves into a new district, other attorneys in that district start to bunch more at 280g. Third, attorneys who I identify as "moved" are often disconnected from their old district in the data and connected to their new district. If the initials-based ID were totally random, we should not expect to see these three patterns.

Date received. The date the criminal case was received by the US Attorney's Office.

Sentence date. For cases that are sentenced, the EOUSA also notes the data of sentencing.

Judge ID. For cases that are brought to a judge, the EOUSA data contains an identifier for the judge involved and that identifier can be linked to a table of judge names. For robustness analyses, I examine the effect of judge race and political party on bunching at 280g. I obtain data on judge characteristics from Crystal Yang's paper on resource constraints and judicial vacancies:

https://test.openicpsr.org/openicpsr/project/114590/version/V1/view?path=/openicpsr/114590/fcr:versions/V1/Data_2015_0150/Public-Use-Data&type=folder