Systemic Disparities in Sentencing: Prosecutors, Race, and Mandatory Minimums*

Spencer Cooper

December 30, 2022

Abstract

I consider how legal institutions may contribute to racial disparities in sentencing outcomes in the US criminal justice system. Specifically, I investigate whether prosecutors bring stronger charges against minorities compared to Whites in federal drug trafficking cases and whether prosecutor choice is driven by the legal framework they operate under. Using a unique data set of US Attorneys and exploiting the bunching point at mandatory minimum drug weights, I show that racial minorities face harsher charges for committing the same crimes as White counterparts and that harsher charges lead to longer sentence lengths. To show that prosecutor charging decisions vary based on legal institutions, I exploit a circuit split in charging law. These laws affect the ability for a prosecutor to bunch defendants. I then use the 2018 First Step Act (FSA) as an exogenous shock to prosecutor incentives and compare effects across circuit law and race. I find that the FSA only reduced racial disparities in circuits where bunching is more costly to prosecutors. This implies racial disparities in charging and sentencing of drug crimes is at least in part driven by systemic factors.

*Affiliation: University of Arizona, Department of Economics. Email: spencercooper@arizona.edu. I thank Evan Taylor for guidance and support throughout this project. I also thank Katherine Barnes, Tiemen Wouterson, Daniel Herbst, Mo Xiao, Peter Hull, Jason Kreag, G. Jack Chin, and several anonymous assistant US Attorneys and Assistant US Attorneys for helpful comments and insight.

It is well documented that Black and Hispanic individuals are disproportionately represented in the United States criminal justice system. A growing literature on crime and race show that racial minorities face discrimination from several legal actors, including law enforcement officers, judges, and prosecutors. However, minorities may experience worse outcomes for committing the same crimes as Whites even if each legal actor they encounter doesn't individually discriminate against them. Rather, a defendant may receive unequal punishment due to a systematic mechanism that creates gaps in outcomes based on some group correlate. In this paper, I show that legal institutions contribute to large racial disparities in charges and sentences of federal drug crimes by impacting the incentives of prosecutors.

Drug trafficking cases are the second most commonly prosecuted crime type at the federal level, making up over a quarter of all cases over the last decade. Nearly all of these cases end with the defendant incarcerated for a long prison spell. In fact, in 2019 drug trafficking cases carried the longest average sentence length for non-violent crime offenses with a mean sentence length of just over 6 and a half years.[1] Thus, drug cases remain a focal point for policy as legal reforms may have large impacts on overcrowded prisons, criminal recidivism, and racial discrimination and disparities.

While the literature and media coverage have mostly focused on racial disparities in mandatory minimum sentencing of crack or cocaine base, large racial disparities exist in mandatory minimum sentencing across all five of the most commonly trafficked drug types. Racial disparities may arise from a number of factors including differences in criminal behavior, differences in mandatory minimum laws across drug type, law enforcement discretion, and judicial discretion. However, a growing body of literature on mandatory minimum sentencing provides evidence that disparities are at least in part driven by prosecutor discretion. Rehavi & Starr (2014) presents striking evidence of this, finding that black defendants receive sentence lengths over nine percent longer than white counterparts and that this disparity is largely driven by prosecutor decisions to file charges carrying mandatory minimum sentences. Similarly, Tuttle (2021) finds that in crack related cases, prosecutors pursue

[1]The only crimes at all with longer sentences than drug trafficking were murder, robbery, kidnapping, and certain sex offenses.

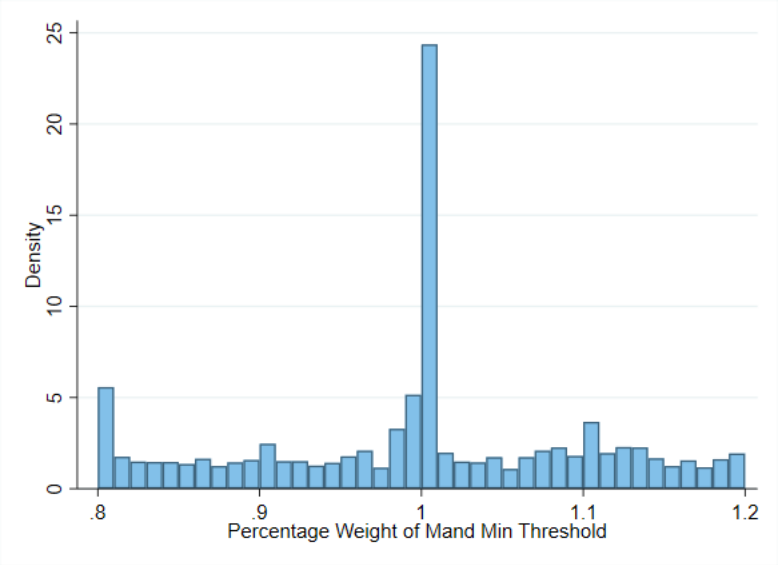
mandatory minimum sentences for black offenders disproportionately more than for white offenders. My findings corroborate this idea, showing that prosecutor decisions drive the race gap in mandatory minimum sentencing. However, I also show that these decisions are significantly impacted the legal framework prosecutors act under, implying the race gap is at least in part driven by systemic factors.

Mandatory Minimum sentences present sentence-floors that a judge must adhere to in their verdict. The cutoffs for whether mandatory minimum sentencing applies are determined by the quantity of drugs for which the offender is charged, measured by the weight. A case that is charged with a weight over the threshold amount will then have a significantly higher expected sentence length, giving the prosecutor significantly more bargaining power in plea deal negotiations. Prosecutors also have the ability to manipulate the charged drug weight to be higher or lower than the weight reported at arrest. Evidence of manipulation is manifested in strong bunching of charging weights at the mandatory minimum threshold amount. Figure 1 illustrates this through a histogram of charged drug weights for weights local to the cutoff and for all drug types. Because the threshold weight is different for each drug type, I normalize weights to be a percentage of the mandatory minimum threshold amount. The graph shows strong bunching right at the cutoff, with small bunching points at other round number percentages.

There is strong heterogeneity by race at this bunching point with racial minorities far more likely to bunched compared to White defendants. Bunching by race is illustrated in Figure 2, which shows that Black defendants are almost twice as likely to be bunched as White ones. This is consistent with Rehavi & Starr (2014) which finds that prosecutors disproportionately charge Black defendants with mandatory minimums, resulting in higher sentence lengths. While differences in bunching illustrated here may be a product of racial heterogeneity in criminal history, conduct, cooperation, or a host of other underlying factors, this disparity still raises concerns of equal treatment and helps motivate this paper's analysis.

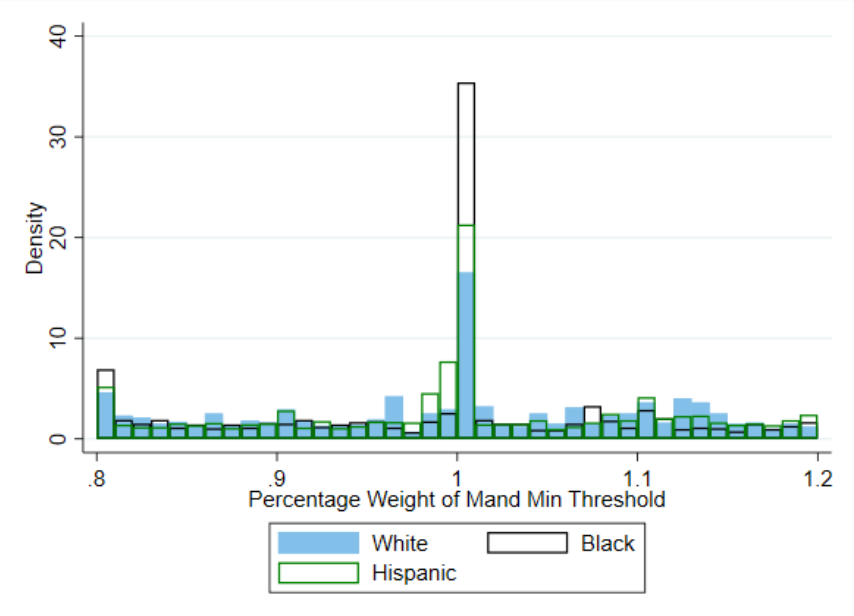
Manipulation and bunching may occur for several reasons including convenience in charging at round numbers, as part of the plea deal bargaining process, or attempts to increase the sentence length. Several papers have postulated that manipulation primarily occurs downward from plea deal negotiations (Bjerk 2017b; Rhodes et al. 2015) while Tuttle (2021)

Figure 1: Charged Drug Weight Distribution: 2013 - 2018



The histogram captures the combined distribution of charged weights across all drug types. Weights are normalized to the high mandatory minimum threshold for comparison.

Figure 2: Charged Drug Weight Distribution by Race: 2013 - 2018



This figure presents three histograms of a similar type shown in Figure 1, one for each race group, stacked on top of each other.

provides evidence of upward manipulation for crack cases. The direction of manipulation is important; if bunching is caused by weights above the threshold being pushed down, it is less likely to have impacts on sentence length and is less indicative of unwarranted racial disparities and discrimination in punitive outcomes. However, if bunching is a product of weights being pushed up into the mandatory minimum range, manipulation may dramatically increase defendants' sentence length and create racial disparities in prison time based on non-legal factors.

Upward manipulation is primarily accomplished by finding evidence that connects the defendant to quantities of drugs outside of the initial seized weight. One of the primary ways prosecutors expand potential evidence is by charging a defendant with conspiracy to traffic drugs (Lynch 2016). Qualifications for a conspiracy charge are broad, with drug conspiracy generally defined as two or more individuals agreeing to transport, manufacture, or sell illegal substances. An individual can only be charged with conspiracy if evidence of such an agreement meets the beyond reasonable doubt burden of proof by which juries operate. Receiving a conspiracy charge often connects a defendant to a larger quantity of drugs or provides more evidence of high drug quantity involvement. In some circuits, defendants convicted of conspiracy can even have their weights added to those of co-conspirators for drugs of the same type.

Thus, the primary cost to the prosecutor to manipulate drug weights are the costs of gathering evidence of additional drug quantities. This is significant to my analysis and to crafting effective anti-discrimination policy for two key reasons. The first is that costs for evidence gathering may vary between racial groups. The second is that evidence gathering costs have been made explicitly lower in certain geographic areas compared to others. There exists a circuit split, caused by varying decisions among circuit courts, in the correct way to consider the charging drug weight for an individual in a conspiracy. In some circuits, prosecutors are allowed to add the drug weights across an entire conspiracy and charge any one member with the full weight. In other circuits, the law specifies that prosecutors must prove additional drugs within the conspiracy are individually attributable to the defendant in question. This effectively has created a set of circuits where drug weight manipulation is significantly less costly for prosecutors than others. I utilize this difference in manipulation

cost to shed light on prosecutor incentives and policy effects by racial group.

In this paper, I answer three main questions relating to racial disparities in federal drug cases. First, I assess whether racial minorities are more likely to be bunched regardless of drug type, defendant observables, or geographic racial composition, and whether this gap is driven by prosecutors. In connection with this, I estimate the effect drug weight manipulation has on sentence length. Second, I consider whether the bunching is a product of weights being manipulated upwards or downwards, further clarifying whether bunching is a source of prosecutors increasing sentence lengths or not. The primary contributions to the literature come from the third question; is the race disparity in charges driven by systemic factors? In conjunction with this, I consider whether decreasing the impacts of mandatory minimum sentences reduces the race gap in charges.

To first identify drug weight manipulation and its impact on sentence length, I proxy for a defendant's probability to be bunched using a residualized bunching propensity measure for the serving US Attorney at the time of the case. The US Attorney acts as the chief federal law enforcement officer in the district and exercises a wide range of discretion in determining which cases are prosecuted and how prosecution procedure is carried out. Thus, US Attorney turnover creates geographic and temporal variation in bunching probabilities. I assume that criminals do not respond to US Attorney service spells, and are thus pseudo-randomly assigned to a high or low-bunching administration. Comparing the bunching propensity measure by race, I find that racial minorities are significantly more likely to be bunched compared to White counterparts. Specifically, a 10 percentage point increase in bunching propensity increases the probability of being bunched for Blacks 3 percentage points more than for Whites. Similarly, I find that high bunching administrations give significantly longer sentences to Black and Hispanic defendants compared to Whites.

After identifying prosecutors as the source of the racial disparity, I then consider whether systemic factors drive the race gap in prosecutors' charging decisions. I utilize the 2018 First Step Act (FSA) as a plausibly exogenous change in prosecutor manipulation incentives and consider the FSA effects by race and across the circuit split. By comparing bunching behavior near the mandatory minimum weight threshold before and after the FSA, I show that the racial disparity is only reduced in circuits where bunching is relatively costly. Fur-

thermore, in circuits with low costs to manipulation, bunching is significantly reduced for White defendants but not for minority ones. Taken together, the results signify that the race disparity is significantly impacted by the legal system prosecutors act under, and that prosecutors consider manipulation costs to be lower for minority defendants compared to White counterparts.

I then provide a series of robustness checks to test the validity of my results. Most importantly, I provide evidence that racial disparities are not driven by differences in criminal behavior but are a result of prosecutor discretion. I compare departures from the sentencing guidelines across race to proxy for differences in observed criminal activity that could impact bunching decisions. I find that departure rates are largely the same across racial groups and that results from both analyses hold when controlling for departure behavior. I likewise provide evidence that bunching is driven from case weights being pushed higher. Finally, I show that geographic variation is best explained by the circuit split in conspiracy law and that other plausible sources of heterogeneity cannot explain the results.

This paper contributes first and foremost to the literature on racial inequality and disparity in the criminal justice system. Racial disparities and discrimination have been investigated and identified across nearly each step of criminal procedure and processing. I add to a growing literature on racial disparities in charging and sentencing decisions (Rehavi & Starr 2014; Robertson et al. 2019; Tuttle 2021; Yang 2016). This paper considers how systemic racial disparities may arise and how policy might mitigate them. Little work has been done in economic literature to estimate systemic factors as a source of disparate outcomes, though Bohren et al. (2022) provide modelling tools for systemic discrimination. This paper especially contributes to the literature focusing on the intersection of race and criminal drug procedure and policy. Findings in this literature vary, with some papers downplaying the significance of mandatory minimums (Bjerk 2005; Fischman & Schanzenbach 2012; Bjerk 2017a; Bjerk 2017b; S. H. Didwania 2020) and others finding significant impacts on sentencing generally and in contributing to racial disparities (Rehavi & Starr 2014; Tuttle 2021).

This paper also contributes to the general literature on prosecutor incentives. This literature has largely focused on identifying what factors prosecutors optimize in charging

decisions. Many of these papers suggest prosecutors maximize sentence length, while others consider conviction rate optimization (Landes 1971; Bebchuk 1984; Glaeser et al. 2000; Gordon & Huber 2002; Boylan 2005; Rasmusen et al. 2009; Silveira 2017). I contribute to this literature by introducing a model that assumes that, rather than maximizing sentence length, prosecutors optimize based on an individual optimal sentence length that varies by defendant. This model allows for trends of prosecutor leniency that have been seen in the literature (Bjerk, 2005).

Methodologically, this paper follows bunching techniques outlined in Kleven (2016) and practiced in Cengiz et al. (2019), Goncalves & Mello (2021), and Tuttle (2021). To construct the bunching propensity measure, I follow the pattern used to construct the judge leniency instrument used in several studies (Kling 2006; Aizer & Doyle 2015; Mueller-Smith 2015; Bhuller et al. 2016; Di Tella & Schargrodsky 2013; Dobbie et al. 2018). However, I use the bunching propensity only as a proxy for bunching probability rather than using it as an instrument.

This paper is one of only a few that illustrates how legal framework can impact prosecutor incentives, how prosecutors respond to a change in incentives, and how prosecutor incentives impact racial disparities in punitive outcomes. It is also the first study to empirically consider the impact and importance of the circuit split in conspiracy law. I show that the circuit split creates different prosecutorial incentives across circuits which are instrumental in crafting effective policy.

More specifically, I contribute to understanding the importance of mandatory minimums by providing evidence that racial minorities are not only more likely to be bunched, but receive higher sentences at the bunching point. This is, to my knowledge, the first paper to empirically estimate the effect of reducing the severity of mandatory minimums on racial disparities in drug sentencing.[2] This paper is also one of the first to consider the impacts of the United States Attorney on case outcomes. While my causal estimates are related to manipulation and overall bunching effects, this study provides evidence that US Attorney administrations vary in prosecution practice and thus create differences in case outcomes.

[2]Tuttle (2021) provides evidence that raising evidentiary standards may reduce bunching but does not provide evidence of how racial disparities may be reduced.

Finally, this is the only paper to provide evidence of drug weight manipulation across each of the five major drug types.

The remainder of the paper is organized as follows. Section II provides background information on mandatory minimum sentencing and drug weight manipulation. Section III presents a conceptual model of prosecutor manipulation choice. Section IV gives details on data. After these three brief sections the paper is split into two major analyses. The first considers disparities in bunching using US Attorney bunching variation in Section V. The second is in Section VI, which considers how systemic factors contribute to disparities and whether policy may be effective in reducing racial disparities. Finally, section VII discusses implications and concludes the paper.

II Background

II. A. Mandatory Minimum Sentencing of Federal Drug Cases

While mandatory minimum sentencing of illegal substances has been practiced in the United States since 1951, the Anti-Drug Abuse Act of 1986 established the punishment framework still used today for federal drug trafficking offenses. It was under this initial framework that drug quantity thresholds were decided, including the infamous 100:1 cocaine to crack ratio. Two threshold limits were set for each drug; a lower threshold applying a five year mandatory minimum sentence upon conviction, and a higher threshold applying ten year sentences. The higher threshold is ten times the amount of the lower threshold for each drug. In order for the mandatory minimum to apply, the charging weight for one drug type must meet or exceed the set threshold weight.[3]

The basic framework established in 1986 largely still applies today, though with some procedural differences. Currently, mandatory minimum sentencing applies to eight illegal or controlled substances: powder cocaine, crack, heroin, marijuana, methamphetamine, lysergic acid diethylamide (LSD), phenylcyclohexyl piperidine (PCP), and fentanyl. Due to few number of cases among the last three types, this paper will focus only on cocaine, crack,

[3]Note that prosecutors cannot sum weights across drug types to get a mandatory minimum charge.

heroin, marijuana, and meth offenses. Offenders still face a lower and higher threshold amount for each drug with a 1:10 weight ratio between thresholds. Table A.1 shows the threshold weights in grams, which are normalized to percentages of high threshold in my analysis. In this paper, I focus only on results surrounding the higher threshold, which has stronger bunching and more severe punishment increases for enhanced cases.[4]

Cases charged at or above the mandatory minimum threshold weight may not necessarily be charged with a mandatory minimum. But being charged at the weight opens the possibility for the prosecutor to impose a mandatory minimum, meaning hitting the threshold weight significantly increases prosecutor bargaining power. For this reason, I do not focus on the mandatory minimum charges themselves in my analysis but only consider the charged weight and the sentence length. Furthermore, mandatory minimum sentences are often non-binding, with many convictions receiving sentences above or below the minimum specified sentence length. Sentences below the mandatory minimum can occur if a defendant provides "substantial assistance" in the prosecution or investigation of a another offender,[5] or if the the offender is eligible for the safety valve provision (described in detail in the Appendix). Bjerk (2017b) notes that over 50% of all cases eligible for mandatory minimum sentencing receive a lower sentence. I likewise find that in my data, only 40% of cases with charging weights over the high weight thresholds receive a sentence at or above the mandatory minimum sentence length.

II. B. Conspiracy Charges and Drug Weight Manipulation

Conspiracy is one of the most common statutes whereby defendants receive mandatory minimums in drug cases. From 2014 to 2020, about 51% of all federal drug trafficking cases carried a conspiracy charge. The conspiracy statute states that "any person who attempts or conspires to commit any offense defined in this subchapter shall be subject to the same penalties as those prescribed for the offense, the commission of which was the object of the attempt or conspiracy."[6] For drug trafficking charges, conspiracy implies a defendant was

[4]Many trafficking cases have a charging weight close to zero. Thus, the lower mandatory minimum weight threshold has a far less prominent bunching point.

[5]See USSG § 5K1.1

[6]21 U.S.C. § 846.

planning or attempting to transport or distribute drugs, often working with others to do so.

Conspiracy charges are also a key method whereby charging weight can be manipulated upwards. For all drug trafficking cases, drug quantity may be determined either from seized quantity or by approximation. Approximation may be based on several factors including cash seized, testimonials, inputs and equipment, laboratory capacity, and estimates of how long drug activity occurred (USSC 2021). This gives the court significant discretion in building a case for the charging drug weight and because of this, drug weight is often a key bargaining factor in plea deal negotiations (Lynch 2016). But in conspiracy cases, prosecutors have even more discretion in drug weight charges. Defendants participating in a drug conspiracy may be charged with weights inclusive of co-conspirators or of the entire conspiracy as a whole. This means that being charged with conspiracy may implicitly carry harsher punishment; the conspiracy statute implies attempting to traffic drugs should be punished the same way as actually trafficking them, and being charged with conspiracy could potentially increase the charging drug weight if prosecutors decide to use manipulation.

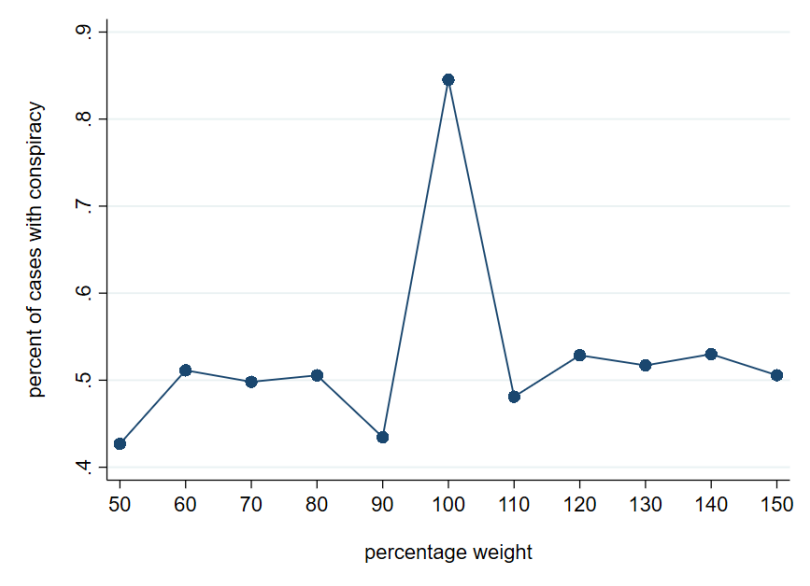
Tuttle (2021) gives evidence of the significance of conspiracy charges. While conspiracy charges are not the focus of the paper, Tuttle shows that the increase in crack weight bunching at the new weight threshold is largely driven by conspiracy cases. In my data set, I also find evidence of a connection between mandatory minimum weight bunching and conspiracy charges. Figure 3 shows the pre-First Step Act percent of cases with a conspiracy charge by charging weight in my primary data set. For local cases below the threshold weight, about 45% carry a conspiracy charge.[7] Likewise, about 47% of local cases above the threshold weight have a conspiracy charge. But at the bunching point, over 83% cases carry a conspiracy charge.

Circuit Split in Conspiracy Case Law

The rules for calculating drug weight in a conspiracy differ across federal circuits, meaning prosecutors face different incentives when seeking to manipulate weights depending on which circuit they are in. The circuits are split between two methods of calculation based on rulings

[7]Here local means within a 50 percent range. So local cases below are all those between 50% the threshold weight up to but not including the threshold weight. Similarly, local above means those above but not including the threshold weight up to 150% the threshold weight.

Figure 3: Percent of cases with a conspiracy charge by weight



Each point gives the average percent of cases with a conspiracy charge for the ten percent bin, besides the bunching point. For example, the point at 50 includes all cases with a drug weight between 50 and 60. The point at 110 contains all cases with weights between 100 and 110, not including the bunching point. The bunching point is all cases with weights exactly at 100.

within each circuit. The first method is the conspiracy-wide, or just conspiracy, approach. The conspiracy approach allows the prosecution to charge every member of a conspiracy with the full drug weight attributed to the entire conspiracy (United States v. Phillips, 349 F.3d 138; United States v. Knight, 342 F.3d 697; United States v. Robinson, 547 F.3d 632; United States v. Stiger, 413 F.3d 1185). The second is referred to as the individual approach. The individual approach holds that each member of the conspiracy must only be charged with the weight "attributable or reasonably foreseeable" to that individual (United States v. Pizarro, 772 F.3d 284; United States v. Rangel, 781 F.3d 736; United States v. Haines, 803 F.3d 713; United States v. Banuelos, 332 F.3d 700; United States v. Stoddard, 892 F.3d 1203). This means that in circuits following the conspiracy approach, prosecutors only need to prove the drug quantity of the entire conspiracy compared to individual approach circuits, where prosecutors must prove quantities attributable to each member of the conspiracy individually.

Because conspiracy charges are a key mechanism for drug weight manipulation, this circuit split creates different costs to the prosecutor for manipulation dependant upon which circuit she practices in. While conspiracy charges in any circuit are likely to increase the

prosecutor's available evidence to connect a defendant to a higher drug weight, doing so is easier in a conspiracy circuit.[8] This is because the prosecutor only needs to prove the defendant is part of a conspiracy and does not need to prove any connection between other conspiracy drug quantities and the individual defendant. Thus, conspiracy approach circuits can be thought of as having a lower cost to manipulation compared to others. Note that circuits where a circuit court decision has not been made are considered high cost circuits. I group this way based on the assumption that circuits without a decision on this matter follow the precedent of *Alleyne v. United States*, which holds that any evidence that increases the mandatory minimum sentence must be evaluated by a jury.[9]

Currently the third, sixth, seventh, and tenth circuits have adopted the conspiracy approach while the first, fourth, fifth, ninth, and DC circuits follow the individual approach. The second, eighth, and eleventh have no precedent at this time. Figure 4 shows these areas and their decisions. Decisions about which approach should be followed are spread across several years, the earliest being in March of 2003,[10] and the latest being in June of 2018.[11] My analyses compare the disparities in bunching and the effect of the FSA on bunching across each method of conspiracy charging law. In both analyses, I find drastic differences in results dependant on circuit type. This highlights the impact incentive structures on prosecutors can have on defendant outcomes.

III Conceptual Framework

To conceptualize the prosecutor incentives and formalize predictions for the FSA policy, I present a simple model of the prosecutor's manipulation choice. I only model upward manipulation, and thus only consider cases with weights below the mandatory minimum threshold weight. Manipulation increases the sentence length for a case but is costly to the

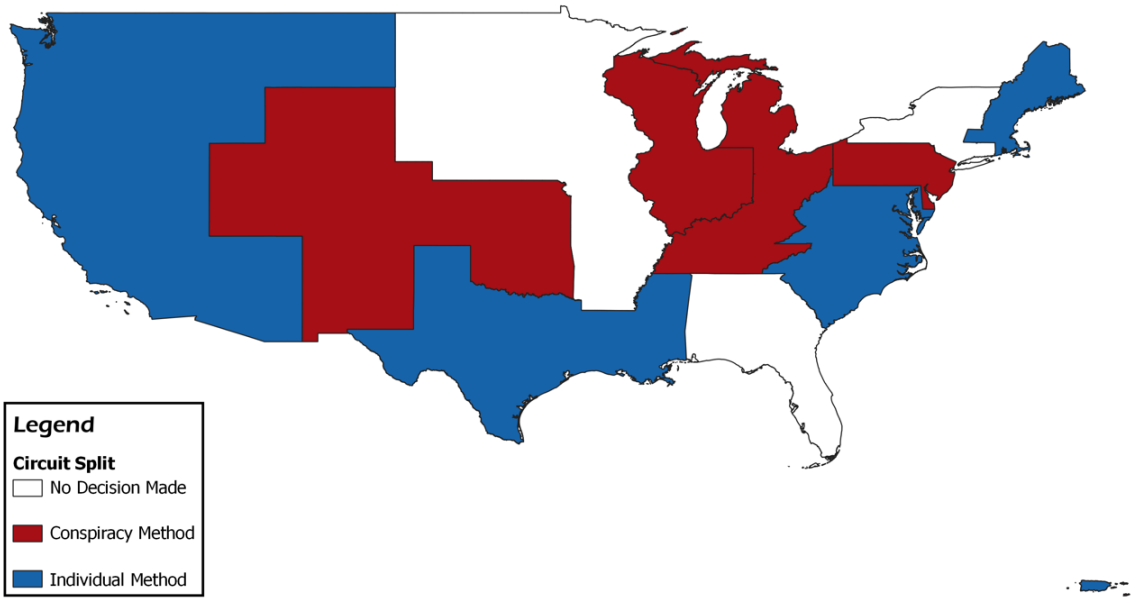
[8] Conspiracy clearly still has an effect on manipulation; when comparing conspiracy charges by weight, the strong bunching seen in Figure 3 holds when only including non-conspiracy approach circuits.

[9] In recent more recent cases, some low cost circuits have mentioned *Alleyne* as potentially conflicting with earlier decisions on conspiracy drug weight methodology. Still, these circuits have not overturned earlier ruling. See *United States v. Gibson* (6th Cir. 2017), *United States v. Young* (6th Cir. 2017), and *United States v. Ellis* (10th Cir. 2017).

[10] *Banuelos*, 332 F.3d 700, (9th Cir. 2003)

[11] *Stoddard*, 892 F.3d 1203, (D.C. Cir. 2018)

Figure 4: Conspiracy Drug Weight Approach by Circuit



prosecutor. Prosecutors will only choose manipulation in cases where a higher sentence will increase utility and outweigh the manipulation costs.

Prosecutors choose a binary manipulation choice m that maximizes their utility function:

$$\max_{m} U(s^*, s, w, \nu) = \pi([s^* - s]^2) - m \cdot C(w; \nu) \tag{1}$$

π gives the utility gained from benefits of sentence length, and is a decreasing function. s^* is the optimal sentence length to the prosecutor for the specific defendant, and reflects a combination of career benefits and what the prosecutor considers fair. Thus, I assume π is maximized when $s^* = s$. C is a cost function that is decreasing in seized weight w. I only model the costs that are added from manipulation and assume all other costs constant regardless of manipulation behavior. ν reflects all other institutional factors that may affect the cost of bunching. s represents the sentence length at the time of disposition. For

simplicity I assume s only takes on three values as described here:

$$s = \begin{cases} s_l & \text{if} \quad m = 0 \\ s_m & \text{if} \quad m = 1 \text{ and } FSA = 1 \\ s_h & \text{if} \quad m = 1 \text{ and } FSA = 0 \end{cases}$$

where $s_l < s_m < s_h$ and $FSA = 0$ if the case happens before the policy change and equals 1 if after.

If the prosecutor chooses not to manipulate weights to the threshold ($m = 0$), they receive the lower sentence length of s_l with no additional costs. If they choose manipulation ($m = 1$), they receive a higher sentence length, s_h or s_m, which may increase utility if s_m or s_h is closer to s^*. They also receive the additional costs of manipulation. Manipulation costs are most likely are made up of costs to search for additional evidence, but may also include increased bargaining costs if the higher sentence increases the time needed to reach a plea deal. C may also include any psychic costs to the prosecutor associated with causing the sentence length to increase.

In my first analysis, I consider variation in bunching behavior by US Attorney spell to identify whether bunching is driven by prosecutor choice. In context of the model, US Attorneys change the utility function π for prosecutors working under their direction. US Attorneys dictate which cases to focus on and how prosecution should be carried out in their district. Thus, for a US Attorney that places high emphasis on drug prosecution the career benefits for the individual prosecutor are likely to be higher than when working for a US Attorney that puts low priority on drug cases.

I assume there are no systematic differences in π by circuit type, meaning prosecutors receive the same benefit of manipulation in each circuit, all else constant. However, conspiracy method circuits are expected to have significantly lower costs to manipulation since evidence gathering costs are explicitly lower. The model suggests that overall bunching should then be higher in conspiracy circuits compared to individual circuits.

Racial disparities in manipulation decisions may come from differences in benefits or costs. Prosecutors may get more benefit from manipulating racial minority charging weights

through some form of taste based discrimination or if bargaining conditions cause sentence lengths to be higher for racial minorities at the threshold weight. Differences in costs could be driven by ease of gathering evidence, unobserved defendant behavior that change bargaining power, different types of conspiracies, or drug carrying behavior being correlated with race.

The First Step Act Predictions

The First Step Act effectively lowered the expected sentence length of mandatory minimum cases. This is reflected in the model as a change from s_h to s_m as the received sentence length for a prosecutor who chooses manipulation. If s^* is sufficiently high, the expected effect of the FSA would be to lower manipulation. As the expected benefits of manipulation decrease, there will be some subset of cases for which the cost will now be too high to justify manipulation. These marginal defendants satisfy two conditions:

1. $U(s^*, s_h, w, \nu) > U(s^*, s_l, w, \nu)$, and
2. $U(s^*, s_m, w, \nu) < U(s^*, s_l, w, \nu)$

These conditions represent a defendant who would have been bunched before the FSA, but no longer are bunched due to the lower benefit. Similarly, the FSA may have no effect on a defendant if they wouldn't have been bunched pre-FSA anyway, or if they would still be bunched after the FSA. That is, there are some defendants who will always be bunched regardless of the FSA, likely because the cost to manipulate their drug weight is extremely low. There are also some defendants who won't be bunched whether their case is before or after the FSA, likely because there is little evidence of a conspiracy or of additional drugs within a conspiracy. Essentially, no change in bunching occurs if manipulation costs are lower than benefits both pre and post-FSA or if manipulation costs are higher than benefits both pre and post-FSA. But bunching may also increase after the FSA. For a given defendant, if costs to manipulation are sufficiently low and s^* is closer to s_l than to s_h but closer to s_m than s_l, the defendant would actually be bunched after the FSA but not before it.

To illustrate how bunching may increase post FSA, I provide a simple example: suppose

for a given defendant and prosecutor, the model parameters are as follows:

$$s^* = 4, \quad s_h = 10, \quad s_m = 5$$
$$s_l = 2, \quad C(w;\nu) = 0$$

Then before the FSA, the distance between optimal and received sentences is lower without manipulation. This means the prosecutor will choose to not manipulate and the defendant will receive a sentence length of 2 years. But after the FSA, the optimal sentence s^* is closer to s_m than to s_l, so manipulation becomes optimal. Then the prosecutor will choose to manipulate after the FSA, but not before. Thus, the effects of the FSA on bunching direction is ambiguous and may vary by circuit and racial group, depending on differences in s^* and $C(w;\nu)$.

The model gives several key insights as to how the results of this analysis should be interpreted. First, decreases in bunching will only happen for groups with sufficiently high costs. Second, groups with high costs but low s^* are likely to have little to no bunching both pre and post-FSA. Third, groups with low cost and high s^* are likely to have high bunching both pre and post-FSA. And finally, increases in bunching will only occur for groups with low costs to bunching and a specific window of s^* that isn't too high or low. These key facts highlight how this analysis uncovers the way prosecutors respond to incentives and how prosecutors may consider manipulation costs between racial groups differentially.

IV Data

I use a unique, hand collected data set of US Attorneys from years 2013 to 2020. The data set includes all US Attorneys from each district, including presidential appointed attorneys, Attorney General appointed attorneys, and acting and interim attorneys that took the leadership role between appointments. This data is used to construct a measure of bunching propensity by US Attorney administration, which is key for identifying racial disparities in drug weight manipulation. This data is gathered from a number of sources including direct correspondence from US attorney district offices, US attorney district office websites,

Wikipedia, and news articles. The data set includes the US attorney's name, nomination date (if applicable), confirmation date, and date out of office. Dates are all recorded at the monthly level to match the drug data, with an attorney being counted as acting that month if the days served are greater than or equal to 16.[12]

Table A.2 gives a few key statistics about US Attorney administrations. For the study period of 2013 to 2020, each district had an average 3.168 attorneys serve in the position with a total of 282 different attorneys. Each attorney served an average of 48.28 months and prosecuted 134.2 drug trafficking cases with weights between 50 percent and 150 percent of the threshold weight. The mean bunching propensity measure is 9.75 but with a standard error of 12.84, indicating high variance between administrations. To further illustrate the variance in bunching, I consider the maximum and minimum bunching propensity measures within each district. The mean maximum propensity is 16.41 and the mean minimum propensity is 2.752.

The primary drug case data is provided by the United States Sentencing Commission (USSC) and includes all federal drug trafficking cases from July of 2013 to September of 2020. I exclude any earlier observations as Supreme Court Case decision Alleyne v. United States may have had an impact on prosecutor behavior regarding drug weight manipulation.[13] Data is restricted to the five most prevalent substances subject to mandatory minimum sentencing: powder cocaine, crack, heroin, methamphetamine, and marijuana.[14] I also restrict the data to the primary racial groups of study: White, Black, and Hispanic. I want to only consider cases local to the threshold weights as weights too high or too low are unlikely to be subject to manipulation at the bunching point. Thus, for the main data set, I only keep cases with weights greater than or equal to 50 percent of the threshold weight and less than or equal to 150 percent of the threshold weight. Because possession cases typically involve lower weights

[12]For some attorneys, dates of entry and exit are only available at the month level. In these rare cases, I default to the incoming attorney being the acting attorney on the month of overlap

[13]Alleyne v. United States, 570 U.S. 99, 108 (2013) holds that "facts that increase the mandatory minimum sentence are therefore elements that must be submitted to the jury and found beyond a reasonable doubt." Effectively, it raised the evidentiary burden of proof for drug weight charging, thereby affecting prosecutor incentives to manipulate drug weights.

[14]The other three substances subject to mandatory minimum law are PCP, LSD, and fentanyl. These had too few observations for any meaningful analysis. There were nearly 3,000 fentanyl cases, but most of these occurred in the year 2019 meaning the difference-in-difference estimates would be unreliable

and are more commonly prosecuted at the state level, I also limit my study to trafficking cases meaning possession only cases are excluded, removing only 26 cases. After excluding any other observations missing key variables, I am left with a data set of 12,722 observations over the seven year period. As a robustness check, I run the bunching analysis with all drug weights. This data set includes 117,432 cases.

The USSC data records each case's charged drug weight in grams, or in gram equivalency of a charged amount (if the case was charged by some non-weight measure). Each observation is a single case and contains details of the defendant's characteristics such as age, sex, race, citizenship, and education; legal details such as month of sentencing, the federal district the case is charged in, the number and types of charges brought forth, and criminal history details; and outcomes such as sentence length and types of punishment. The charging details also allow me to distinguish between cases that have received a conspiracy charge and those that have not.

Table A.3 provides summary statistics for the USSC data set. I provide statistics for the entire sample as well as for the different circuit types individually. Individual approach circuits make up just over three quarters of the sample of cases. This is due to only four circuits with case law precedent for low costs to manipulation and the fact that three of these circuits have relatively low caseloads. Panel A. gives the distribution of drug type charges. Note that crack cases, the primary focus of the literature, contains less than 11% of all charges for this sample. Proportions of cases for each drug type are fairly similar across circuit groups. Panel B. gives statistics for defendant characteristics. Nearly half of the sample is made up of Hispanic defendants and 84.8 percent of all defendants in the studied drug weight range are male. Panel C gives figures for outcome variables and shows that around half of all cases in the data set have a conspiracy charge. The mean sentence for these cases is 75 months, signifying the high punishment level for these types of drug crimes. Across defendant characteristics and outcomes, figures are very close between circuit cost groups. The main discrepancy is in racial makeup, where low cost circuits have less Hispanic cases compared to high cost ones.

In addition to the US Attorney and USSC data, I use census data to control for racial population changes over time at the state level. I also use monthly prison admission and

population data provided by the National Corrections Reporting Program made available from the Bureau of Justice Statistics. This data measures prison statistics for state level convictions by crime type. The prison data is used to see if disparities are driven by differences in state-level drug punishment opportunity.[15] I also use google trends search data from Stephens-Davidowitz (2014) to consider how my results correlate geographically with racial animus. Finally, I collect statistics on the 2016 presidential election results by state from the New York Times. These are used to see if disparities in manipulation can be explained by political differences across states.

V Identifying Bunching Disparities

The goal of this section is to identify racial disparities in drug weight manipulation and show they are a result of prosecutor discretion. Disparities in bunching may arise from a host of different sources and may not be indicative of unequal treatment. However, if the bunching at the threshold weight is the result of prosecutors disproportionately targeting minorities holding other legal factors constant, then bunching represents a systemic or discriminatory disadvantage for racial minorities. To identify prosecutor effects, I consider differences in bunching behavior between United States Attorney spells.

V. A. United States Attorneys

United States Attorneys serve as the chief federal law enforcement agent within their district. There are 93 US Attorneys in the US at all times, one for each district. US Attorneys are typically appointed by the president of the United States and serve until they choose to step down or are asked to resign. Resignation requests often occur after a new president is sworn in, but may also occur within presidential administrations. In times of vacancy, an assistant attorney already serving in that district fills the leadership role and is considered the Acting US Attorney. For extended vacancies, the US Attorney General may also appoint an interim attorney to fill the leadership role until a new presidential appointment.

[15] I utilized this data from ICPSR here: `https://www.icpsr.umich.edu/web/NACJD/studies/38048`

Attorneys are given immense discretion to dictate the focus and procedure of prosecution within their district. In the Principles of Prosecution section of the United States Justice Manual, it states that "...individual United States Attorneys are required to establish their own priorities (in consultation with law enforcement authorities), within the national priorities, in order to concentrate their resources on problems of particular local or regional significance."[16] Other sections of the Justice Manual describe the US Attorney as having "the broadest discretion in the exercise of such authority" in relation to prosecuting criminal matters.[17] This implies US Attorneys have significant impacts on the types of cases that are prosecuted and the manner in which prosecution should occur. This appears to hold true in practice; in my discussion with a number of federal prosecutors, they described the US Attorney in office as having a significant impact in the day to day operations of the Assistant US Attorneys.[18]

While individual prosecutors certainly have their own discretion, general trends in prosecution may be better explained at the US Attorney level.[19] Specifically, I consider how bunching may vary by US Attorney service. A US Attorney that focuses on harsh prosecution of drug cases may lead to an increase in bunching during their term compared to other districts at that time and compared to other US Attorneys within that district at other time periods. In support of this, I find strong variation in bunching by US Attorney spell. As shown in Table A.2, the US Attorney bunching propensity measure has a high standard deviation of 12.84. I also calculate the standard deviation within districts to ensure variation is not only driven by geographic differences. The within district standard deviation is 7.45, a number still high relative to the overall mean. I also show that variation is not driven only by temporal differences; Figure A.2 displays bunching for each district averaged across each US Attorney in the district. The figure shows high variation in bunching across districts.

[16] See section 9-27.230.

[17] See section 9-2.001. This section also gives specifics about for which aspects of prosecution US Attorneys are allowed to exercise their discretion. This includes authorizing prosecution and determining the manner of prosecuting and deciding trial related questions.

[18] I spoke with a handful of US Attorneys and Assistant US Attorneys about the impacts of who is in the role of US Attorney, who are kept anonymous by request. While responses varied, the general consensus is that US Attorneys have a lot of flexibility in what types of cases should be prosecuted and how prosecution should be carried out. Conversations with legal scholars have told a similar story.

[19] I do not have individual federal prosecutor data so I cannot check to what level individual prosecutors bunch compared with US Attorney spells.

Finally, Figure A.3 illustrates variation within district by taking the difference between the max and min bunching propensity measures for attorneys within a single district. Each of these figures suggest high variation in bunching across US Attorney spell.

V. B. Empirical Strategy

I utilize the variation in bunching propensity by US Attorney administration as a pseudo-random measure of the probability of being assigned bunching. The idea is that non-bunching and low-bunching administrations serve as reliable counterfactual distributions compared against high-bunching administrations. This is a similar strategy as described in Frandsen (2017) and practiced in Goncalves & Mello (2021). In their analysis, Goncalves & Mello (2021) estimate a binary measure of bunching and compare bunching versus non-bunching officers. Rather than trying to determine a cutoff between bunching and non-bunching administrations, I use a residualized, continuous measure of bunching propensity to identify prosecutor-driven racial disparities. This approach has a flavor of judge-leniency instrumental variables as seen in a large number of law and economics papers (Kling 2006; Aizer & Doyle 2015; Mueller-Smith 2015; Bhuller et al. 2016; Di Tella & Schargrodsky 2013; Dobbie et al. 2018). However, I don't use the propensity measure as an instrument but consider it a proxy for the defendant's probability of being charged at the bunching weight.

To estimate racial disparities, I employ the following model:

$$\begin{aligned} Y_{idmt} = \alpha + \beta_1 Black_i + \beta_2 Hispanic_i + \beta_3 bunch\_score_{idmt} \\ + \beta_4 Black_i \times bunch\_score_{idmt} + \beta_5 Hispanic_i \times bunch\_score_{idmt} \\ + X_i\gamma + P_{dt}\delta + \lambda_d + \kappa_t + \eta_m + \epsilon_{idmt} \quad (2) \end{aligned}$$

with β_4 and β_5 being the primary coefficients of interest. These give the differential effects of increasing bunching propensity on Black and Hispanic defendants compared to White counterparts. X_i gives observable characteristics of the case and of the defendant. These include the drug type, the total number of criminal history points applied in the case, the sex of the defendant, defendant age and age squared, a binary for whether the defendant

completed some or graduated from college, and a binary for whether the defendant is an illegal alien. It also contains a measure of the percent of cases that are White per US Attorney spell to control for differences in case composition and prosecution opportunity across race. P_{dt} is a set of annual race population proportion measures to control for differences in racial make-up of the district. I likewise include district, year, and month-of-year fixed effects. The fixed effects control for seasonality, temporal shocks, and district differences in bunching behavior.

To construct the bunching propensity measure, I use a leave-out, residualized mean of US Attorney bunching decisions, similar to the approach used in Dahl et al. (2014), Dobbie et al. (2018), and Arnold et al. (2018). I regress the bunching decision on year-by-district fixed effects and month of year-by-district fixed effects. I then take a leave-one-out mean of the residuals across each US Attorney, where each individual's bunch score value is a mean across all other cases for the US Attorney in office during the individual's case timing and location. Controlling for time-by-location fixed effects is necessary since US Attorney assignment is not truly random. High or low bunching attorneys may be assigned to a district at a specific time in response to drug crime behavior in the area.

The primary assumption of this method is conditional random assignment of bunching propensity to a defendant. Another way of saying this is that whether a defendant is prosecuted by a high bunching or low bunching administration is effectively random after controlling for district-by-time fixed effects. This implies individuals are not changing their criminal behavior based on the current or recent US Attorney prosecution patterns. This seems reasonable; to have a meaningful impact on criminal behavior, criminals would need to be aware of the US Attorney's position on drug prosecution and that knowledge would need to have a strong enough incentive to change production or transportation activity. Beyond this, many federal cases pass through the state system first, meaning there is uncertainty for the defendant about which level they will be prosecuted at. A bigger threat comes by way of detection and law enforcement activity. US Attorneys work in close contact with members of the FBI and sometimes with the US Marshall's Service or the DEA. A high bunching attorney may also encourage specific types of drugs be targeted or more arrests in general.

While I cannot test directly for changes in criminal or law enforcement, I can broadly test

for selection on observables by regressing bunching propensity on defendant characteristics. Table A.4 gives F-Statistics and tests for regressions run on observables. The first column tests just for bunching just as a binary variable. The joint F-test for this specification returns an F-value of 4.74 and a p-value of 0.000. The next three columns then regress bunching propensity on observables. This shows that the F-value decreases significantly from the bunching decision to the bunching propensity measure, indicating conditional random assignment. The p-value also increases and is insignificant across each tested sample. To test changes in law enforcement, I consider the main specifications again but drop attorney spells with especially low or high cases per month. This is discussed in more detail below in the robustness checks section of the results.

Finally, it may be that racial disparity is driven by unobserved characteristics of the case that are correlated with race that change the prosecutors bargaining power. For instance, if racial minorities happened to sell to minors more often, that may increase the expected sentence length but would be unobservable in the data. Randomization by attorney propensity may help control for this, but it is possible that low bunching attorneys ignore these effects while high bunching ones exploit them. I argue the racial disparity is not driven by such factors by comparing departure rates between race groups. Departure means an individual is charged outside of the recommended sentencing guidelines for some specific reason. A departure can increase the sentence if the defendant has what are called aggravating factors; these include things like cruelty, involving a minor, or having a leadership role. Departure can decrease a sentence if the defendant has mitigating factors; these include things like having a minimal role in selling or being a minor. Thus, if unobserved case characteristics were driving the bunching disparity, the rate of departure would be different by race. I find this is not the case. Furthermore, I find that controlling for departure does not effect the statistical significance of the results nor does it substantially affect magnitudes of the estimates.

V. C. Results

Table 1 displays the main estimates for heterogeneous bunching by race. The first three columns control for state differences in racial population proportions by year, year and month fixed effects, district fixed effects, and the timing of the FSA policy. They do not include

any of the defendant characteristics. This is because disparities could be caused by some other discriminatory factor that correlates with race. This is an important distinction; if this is the case, racial disparities in bunching likely still indicate worse outcomes (through longer sentences) for racial minorities but the source is not in the defendant race necessarily. For example, prosecutors may bunch individuals with higher criminal history records more than low history offenders, a variable that correlates with race. In columns 4, 5, and 6, I include the defendant characteristics in addition to the population controls and fixed effects. For each level of controls, I estimate the effects for the full sample, the sample excluding attorneys with less than 25 total cases, and the sample excluding attorneys with less than 50 total cases.[20] I reduce the sample this way because attorneys with few cases may have extremely large or small propensity measures simply because of the small case count. Thus, I compare results across specifications to make sure outliers are not driving results.

The estimates indicate Black defendants face significantly higher bunching odds when assigned to a high bunching administration. Specifically, a 10 percentage point increase in bunching propensity of the US Attorney administration increases the probability of a Black defendant getting bunched by 3.33 to 4.26 percentage points over White counterparts. This represents a 34 percent to 43.6 percent change compared to the overall bunching mean. These estimates are statistically significant at the 5 percent level across the specifications. Consistency in magnitude and significance of estimates regardless of the inclusion of defendant characteristics indicates that observable characteristics are not driving the results. Effects for Hispanic defendants are also positive, indicating higher bunching for them over Whites, but the effects are smaller and only marginally significant dependant on specification.

I also test for whether higher bunching attorneys assign higher sentence lengths to racial minorities. I use the same method outlined in Equation 2 with sentence length as the dependant variable. Note that this is not necessarily a causal estimate of how bunching impacts sentence; to estimate that I would need to use US Attorney bunching propensity as an instrument for being bunched, similar to how judge leniency is often used to instrument for incarceration or sentence length. I refrain from using the propensity measure as an

[20]Note that the number of cases mentioned here is just the number of drug trafficking cases within the 50 percent - 150 percent weight window. Even US Attorneys with relatively short spells see many cases when including all weights and other types of crime.

Table 1: Racial Disparity in Bunching

	(1)	(2)	(3)	(4)	(5)	(6)
Black	-0.00180	-0.00166	-0.00767	-0.0192	-0.0180	-0.0227
	(0.00902)	(0.0103)	(0.0122)	(0.0117)	(0.0134)	(0.0155)
Hispanic	0.00605	0.0113	0.00555	-0.0211**	-0.0165*	-0.0180
	(0.00828)	(0.00903)	(0.00980)	(0.00914)	(0.00969)	(0.0111)
bunch_score	0.00631***	0.00756***	0.00800***	0.00616***	0.00745***	0.00750***
	(0.000931)	(0.00107)	(0.00161)	(0.000928)	(0.00116)	(0.00178)
Black*bunch_score	0.00358**	0.00332**	0.00417**	0.00356***	0.00334**	0.00426**
	(0.00139)	(0.00155)	(0.00203)	(0.00135)	(0.00149)	(0.00196)
Hispanic*bunch_score	0.00182**	0.000963	0.00129	0.00164*	0.000813	0.00120
	(0.000915)	(0.000954)	(0.00132)	(0.000857)	(0.000957)	(0.00131)
Covariates				X	X	X
Population Controls	X	X	X	X	X	X
Fixed Effects	X	X	X	X	X	X
Attorney # of Cases	All	≥ 25	≥ 50	All	≥ 25	≥ 50
R Squared	0.182	0.187	0.180	0.200	0.205	0.194
N	12,216	10,927	8,553	12,216	10,927	8,553

Notes: The dependant variable is a binary variable for whether the charging weight is at the threshold weight (bunching point). The first three specifications omit dependant covariates to assess whether racial minorities face more bunching while allowing for correlations with observables in the measure. Specifications 4-6 then control for these, comparing individuals of similar drug types, criminal history, and other observable traits. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

instrument because a US Attorney who bunches more may also affect many other parts of the sentence length, implying a violation of the exclusion restriction. However, this analysis is still informative and speaks to the importance and impact of mandatory minimum charges. If bunching is associated with increases in sentence length, then racial disparities at the bunching point indicate harsher punishment for minorities.

I first present the raw averages in sentence length by weight and race in Figure 5. Each point represents a mean of the ten percent weight bin, besides at the bunching point which is just the individual bunching point weight. This gives approximately equal number of cases for each point. Given all weights above the bunching point are subject to the 10 year mandatory minimum and all weights below are not, one might assume relatively flat trends on either side of the threshold weight with a discontinuous increase the the threshold weight. This is largely true for White defendants. However, for racial minorities there is a bunching increase in sentence length at the weight bunching point while there is essentially no effect

Figure 5: Sentence Length By Charging Weight

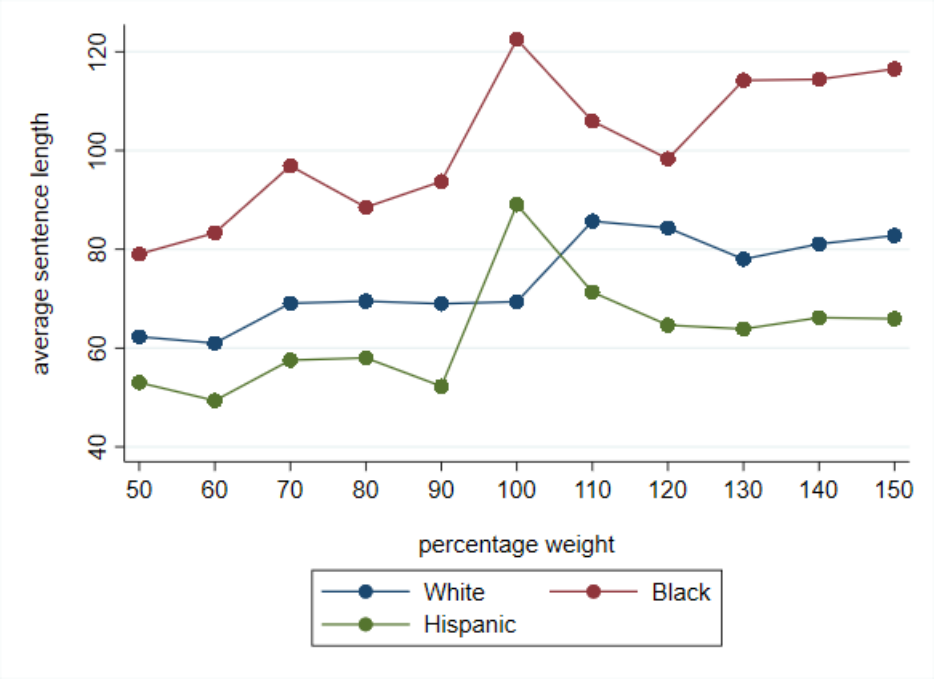


Figure presents race specific mean sentence lengths for charging weight bins of 10 percent, besides at the bunching point which is measured alone. For cases below the bunching point, bins are rounded. For example the 50 point contains all cases with weights at or above 50 percent but below 60 percent. Points above the bunching point are rounded down. This separates the sample into approximately equal bins.

for Whites. If this sentence length bunching is driven by drug weight manipulation cases, it indicates one of two things about prosecutor behavior. First, bunching minorities may be more beneficial to prosecutors than bunching Whites. Second, bunching minorities may cost less than bunching for Whites does. Either way, if sentence length bunching is caused by drug weight manipulation it is indicative of disproportionate punishment for similar types of defendants.

The regression analysis tests whether sentence length disparities are driven by drug weight manipulation. Table A.5 displays the estimates for racial differences in sentence length associated with bunching propensity. The first three specifications give the main results and show significant increases in sentence length for both Black and Hispanic defendants compared to White defendants. The estimates suggest a 10 percentage point increase in bunching propensity is associated with a 2.47 to 3.85 month increase in sentence length for Black defendants and a 2.21 to 2.94 month increase in sentence length for Hispanic defendants compared to White counterparts. This represents a 3.28 to 5.11 percent increase

for Blacks and a 2.93 to 3.9 percent increase for Hispanics over the average sentence length. Specifications 4-6 run these same regressions but omit the bunching point from the sample. This highlights whether differences in sentence length are coming from manipulation to the bunching point or not. I find small, insignificant effects across each of these specifications, suggesting sentence disparities are driven by drug weight manipulation and not US Attorney spell effects at other weights.

Robustness Checks

To check that results are not driven by law enforcement changes or substantial changes in volume of cases prosecuted, I run the above analysis again omitting attorney spells with especially high or low cases per month. Table A.7 displays these results. Columns 1 and 2 omit US Attorney spell cases with the bottom 10 percent of cases per month while columns 3 and 4 omit the top 10 percent. These estimates have similar magnitudes to the main specifications and remain statistically significant.

I also consider the main results again using a binary framework of treatment rather than a continuous measure. This is closer to the approach used in Goncalves & Mello (2021), where the estimates are traditional difference-in-difference results comparing non-bunching to bunching groups across race. I distinguish bunching from non-bunching attorneys by comparing the percent of cases charged at 99 percent weight, 101 percent, and at the bunching point. I take a ratio of the bunching weight compared to all weights in this window. The US Attorney bunching treatment equals one if this ratio is strictly greater than one third.[21]

The results of the binary analysis are presented in Table A.8. They are consistent with the main analysis above. I find bunching attorneys increase bunching for Blacks between 1.9 and 3.56 percentage points over White counterparts. I also find Hispanics more likely to be bunched compared to Whites using this method, though the results do not survive the inclusion of covariates. This implies Hispanics are receiving more bunching but through some observable factor correlated with race. Statistical significance for these estimates is marginal and varies by specification. Similarly, in the binary analysis for sentence length I find racial

[21]This is essentially a simplification of the method described in Frandsen (2017) and used in Goncalves & Mello (2021). The main difference is I do not estimate a distribution curvature parameter k or test for the probability an attorney falls within the estimated range.

minorities disadvantaged, with higher sentences associated with bunching attorney spells. Once again the results here are less statistically significant than when using the continuous measure.

VI Systemic Impacts on Disparities

The analysis in the previous section gives evidence that racial disparities in charging weight bunching are driven by prosecutor drug weight manipulation choices, and that these choices have significant impacts on sentence length. This section considers how institutional factors impact the disparity and how policy might reduce it. Exploiting the circuit split in conspiracy law and the FSA policy adoption, I show that Blacks and Hispanics are systematically more likely to be bunched than their White counterparts. I first discuss the FSA policy in detail, then present the empirical strategy for assessing the impacts of the policy across circuit groups. The results shed further light on how prosecutor's incentive to manipulate varies based on race.

VI. A. The First Step Act

The First Step Act of 2018 (P.L. 115-391) was signed into law on December 21, 2018. The law has five major titles, which can broadly be summarized as recidivism reform, prison firearm storage security, treatment of pregnant prisoners, sentencing reform, and a series of other miscellaneous criminal justice provisions. Here I focus only on the description of sentencing reform as the other titles are likely irrelevant for this study. The two major and most relevant sentencing reforms involve reducing and restricting sentence enhancement for prior drug felonies, and expanding what is known as the safety valve provision, whereby defendants with low criminal history and non-violent charges can be exempt from a mandatory minimum sentence.[22]

Prior to the FSA, having a charging weight at or above the high threshold not only

[22]The sentencing reforms involve four major revisions to the law. The other two not listed above involve reducing the severity of "stacked" § 924(c) offenses, and retroactively applying the Fair Sentencing Act of 2010. These provisions should have little to no effect on drug weight manipulation incentives for prosecutors. See `https://www.congress.gov/bill/115th-congress/senate-bill/756` for full details of the law.

secures a 10 year mandatory minimum sentence upon conviction, but also opens the door for a 20 year enhancement. This enhancement is specified based on the defendant's prior criminal history. These punishments are significantly higher than those for weights just under the threshold, where the statutory mandatory minimums are five years and 10 years with enhancement. Furthermore, before the FSA if an offender had more than one criminal history point, they were not eligible for the safety valve provision, which effectively removes the mandatory minimum. The FSA decreased the punishment for enhanced offenders at this high threshold, decreased the pool of individuals who can receive enhancement, and increased the pool of individuals who qualify for safety valve. Taken together, the sentencing revisions from the FSA effectively lowered the expected payoff of charging at the threshold weight. I provide a more detailed explanation of enhancements and the safety valve provision in the Appendix.

VI. B. Empirical Strategy

To compare effects differentially by race, I employ a difference-in-differences approach to the bunching analysis:

$$Y_{idtma} = \alpha + \beta_1 minority_i + \beta_2 FSA_{tm} + \beta_3 minority_i \cdot FSA_{tm} + X_i\gamma + P_{dt}\delta + \tau_{dtm} + \lambda_d + \delta_m + \omega_a + \epsilon_{idtma} \quad (3)$$

Here Y_{idtma} measures the outcome of interest for case i in district d during year t and month m, prosecuted under US Attorney administration a. Outcomes include the binary bunching measure, the binary measures of just above or just below the threshold, and a binary measure for whether case i included any conspiracy charges. FSA_{tm} is a binary measure for post-FSA timing, December of 2018, which is used to capture the effects of a change in prosecutor incentives. The set of controls for this model is similar to those used in the attorney bunching model, specified in Equation 2. The main differences are that I exclude year fixed effects to avoid collinearity issues since the post-FSA period is so short, I include US Attorney fixed effects, and I include a set of circuit court decision binary variables, represented as τ_{dtm}. The circuit court decision variables account for the timing and geography in conspiracy charging

decisions.[23] I also exclude the percent of cases that are White by attorney spell measure, which is now captured by the US Attorney fixed effects. β_2 and β_3 are the primary coefficients of interest. β_2 gives the effect of the FSA on manipulation behavior for White defendants. $\beta_2 + \beta_3$ then gives the effects for minority race defendants.

Each of the dependent variables are binary measures for cases falling within a specific weight range. I define a case as bunched if the charging weight is exactly equal to the threshold weight. In my robustness checks, I vary this definition to allow for some cases just past the threshold weight.[24] To assess rounding directions I similarly consider cases that are "just above" or "just below" the threshold. It is not clear to what degree prosecutors are able to manipulate weights and thus how large to make these "just above" and "just below" windows, motivating the need for many checks with different windows of weights.

The basic strategy for assessing rounding direction is this: given a decrease in bunching occurs for a some racial-circuit group, the missing mass from the bunching point will appear as excess mass in some other part of the distribution. Under the assumption of the conceptual model, the excess mass signifies the prosecutor is now charging at the seized weight. This means the excess mass describes whether weights are being manipulated upward or downward. If cases are experiencing upward manipulation, that is, prosecutors push their charging weight higher to get a higher sentence, the excess mass will appear below the bunching threshold. I assume the cost of manipulation decreases with seized weight, implying that cases near the threshold but just below it are more likely to be subject to manipulation. Then if upward manipulation occurs, the prediction would be an increase in excess mass near the bunching point but to the left side of it. If bunching is a product of rounding down, the excess mass would appear above the threshold weight. This idea is illustrated in Figure A.1.

The main threat to identification in these models is systematic or behavioral differences across racial groups. Case types may vary across racial groups for several reason. One key

[23]These are simply binary measures for a circuit court decision to adopt a specific conspiracy weight charging approach. For example, the 5th circuit decision to adopt the individual approach occurred in October of 2015. Then the binary measure for that decision only equals one for cases in circuit 5 that occur after October 2015.

[24]I use a bunching measure that includes weights greater than or equal to the threshold weight and less than or equal to 1.02 times the threshold weight. I use this window approach to account for differences in magnitudes of threshold weights which may lead to differences in manipulation behavior across drug types.

difference may be in the distribution of actual drugs carried. Another may be in enhancement type behavior that increases expected sentence length beyond the primary charges. As discussed in Section V, this does not appear to be the case. Another concern would be if racial minorities were more likely to be involved in conspiracies or were in different types of conspiracies, such as gang involvement. Conspiracy charges and drug weight manipulation are also often products of plea bargain negotiations and are significantly affected by the defendant's propensity to assist in the prosecution of others. Differences in the propensity to assist the government across races may therefore also confound the effects of the above models.

Issues of differences in carrying distributions are largely mitigated due to the narrow window of weight I look at which are local to the mandatory minimum threshold. The weight cutoffs for the sample are arbitrary so I consider other cutoffs, including the full distribution of weights, in the robustness section below. Gang affiliation data is, to my knowledge, inaccessible meaning controlling for gang activity is nearly impossible. One indication of this may be conspiracy charges. Individuals with large, easily proved gang affiliations may be more likely to receive conspiracy charges. However, when considering the full distribution of weights, Whites have the highest percentage of conspiracy with 58 percent of cases having at least one conspiracy charge. When looking at the sample window, Black defendants have the highest percentage at about 57 percent, which is about ten percentage points higher than for Whites and Hispanic defendants. I also try to mitigate gang activity effects by dropping districts with the highest level of influence of major Mexican Transnational Criminal Organizations according to the DEA. Finally, I find that Black defendants do receive less government sponsored sentence reduction than Whites or Hispanics, implying a lower willingness to cooperate. However, the difference is relatively small with around 37 percent of White cases, 35 percent of Hispanic cases, and 30 percent of Black cases receiving a government sponsored downward departure for cases in the main data set. As stated in section V. B., departure rates relating to case characteristics are nearly identical between race groups.

One other concern is that conspiracy circuit decisions may not be actually driving the heterogeneity, but some other correlated factor. While I cannot test for this directly, I

provide a series of alternate explanations using other plausible mechanisms in the robustness section below.

VI. C. Results

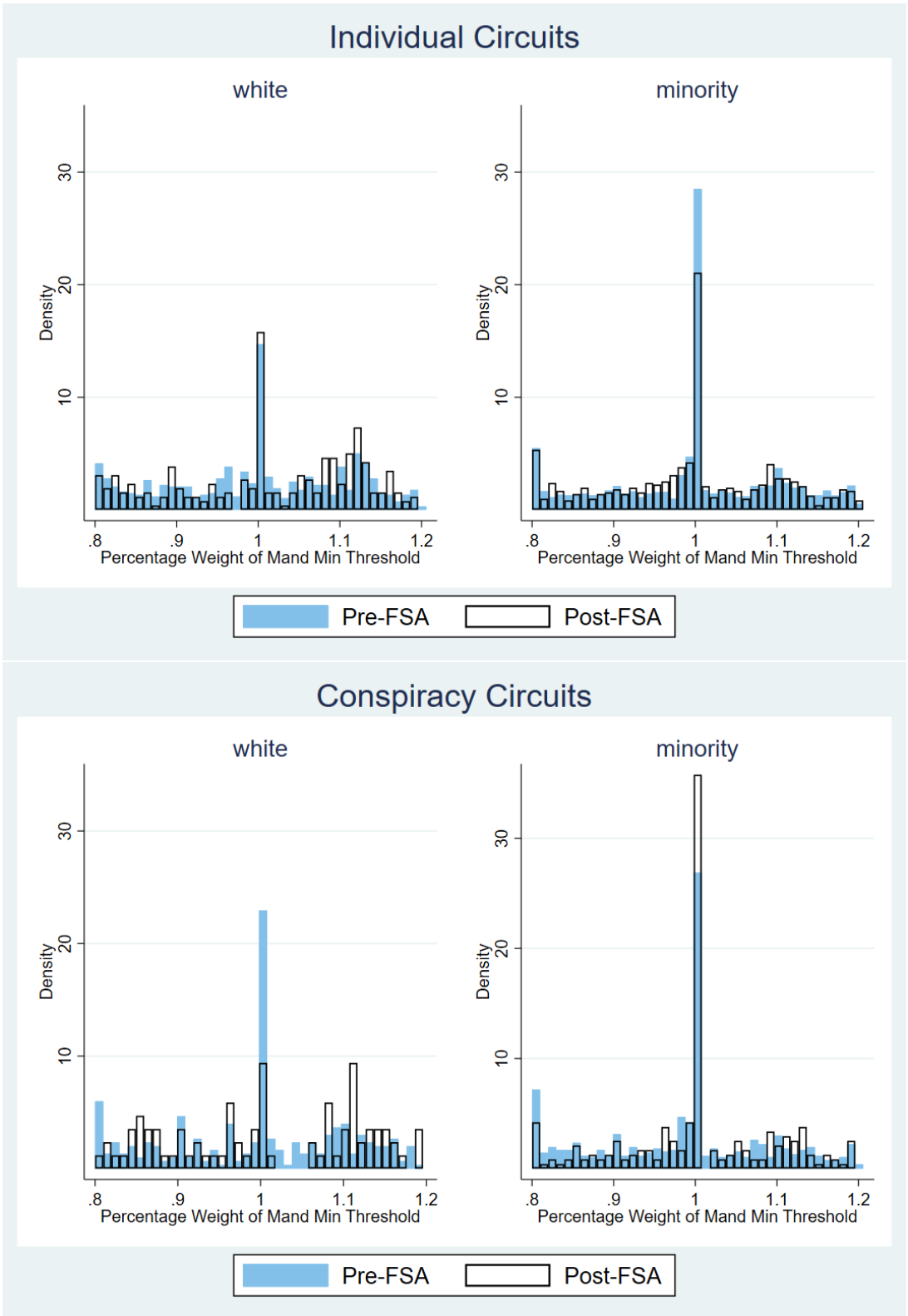
The FSA caused a large, immediate drop in bunching at the high threshold. Figure A.5 shows a dramatic decrease in cases charged at the mandatory minimum threshold weight. As illustrated in the model in Section III, a decrease in bunching implies prosecutors reduced drug weight manipulation for some set of marginal defendants. The model also suggests that differences in cost C or optimal sentence length s^* between defendant groups should cause heterogeneous responses to the FSA. To test this, I compare the effects of the FSA across conspiracy and individual circuits and across White and minority race defendants.[25] Groups with a significant decrease in bunching must include cases with higher costs C. Groups with increases likely contain defendants for whom the costs to manipulation are low. Groups with no change in bunching either had little bunching before hand or are comprised of cases that get bunched before and after the FSA. The model suggests these cases should have relatively low costs to manipulation. Thus, I use the bunching responses to back out how prosecutors are incentivized to manipulate drug weights differentially across race.

Figure 6 displays the heterogeneous effects of the FSA on bunching. In circuits where drug weight manipulation is more costly, the FSA caused a large decrease in bunching for minority defendants and almost no change for White defendants. Note that this means the FSA reduced the race gap in individual method circuits. However, in conspiracy method circuits the racial disparity actually increases. This is because when costs to manipulation were high, large decreases in bunching occurred for White defendants while minority cases experienced almost no change.

Table 2 displays the main regression results. Columns 1 and 2 display the results for individual circuits while 3 and 4 show the results for conspiracy circuits. For each set of circuits, I include one specification controlling only for defendant characteristics and circuit

[25]Note, in this section I display estimates for Black and Hispanic defendants together. This is because the effects for each race group is almost identical in both the high cost and low cost circuits. See Table A.6 for Black and Hispanic effects separately.

Figure 6: Effects of FSA by Circuit Group and Race



decision controls, and another specification that adds racial population controls and month of year, district, and US Attorney fixed effects. The regression results corroborate the effects illustrated in Figure 6. For both specifications, in the individual circuits the FSA had a strong, statistically significant effect on minority defendants. Minority case bunching decreased between 3.65 and 5.29 percentage points compared to White counterparts following the FSA, a 36.7 to 53.2 percent change from the pre-FSA mean. White cases experienced no significant change in bunching. In conspiracy circuits the FSA significantly reduced bunching for White cases, with a decrease between 6.67 and 8.83 percentage points, a 61.3 to 81.2 percent change from pre-FSA mean. The coefficient for minority case differential bunching ranges from 9.69 to 8.99. This means minority case bunching actually increases post-FSA.

Table 2: Changes in Bunching by Race and Circuit Type

	Individual Circuits		**Conspiracy Circuits**	
	(1)	(2)	(3)	(4)
minority	0.0207*	0.0154	-0.0580**	-0.0460**
	(0.0122)	(0.0113)	(0.0212)	(0.0196)
fsa	0.00955	0.0130	-0.0667***	-0.0883***
	(0.0156)	(0.0153)	(0.0225)	(0.0310)
minority*fsa	-0.0365**	-0.0529***	0.0969***	0.0899**
	(0.0146)	(0.0142)	(0.0240)	(0.0327)
Covariates	X	X	X	X
Circuit Decision Controls	X	X	X	X
Population Controls		X		X
Fixed Effects		X		X
R Squared	0.066	0.226	0.061	0.165
N	9,643	9,637	3,079	3,079

Standard errors in parentheses

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

Notes: The dependant variable is a binary variable that equals one if the charging weight is exactly equal to the mandatory minimum threshold weight.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure A.6 displays these main results across time by plotting the difference-in-bunching coefficients by bi-annual intervals. These figures also serve to check that pre-trends are parallel. I do this separately for individual and conspiracy circuits. Both figures show flat pre-trends. Figure A.6a shows the immediate decrease in bunching for minorities in

individual circuits compared to white defendants. This effect appears to potentially return to pre-trend levels over time, something that will be further investigated as current data becomes available and implemented. Figure A.6b shows a potential anticipation effect in conspiracy circuits, where bunching increases just before the FSA passage and remains high afterwords.

Direction of Manipulation

From a welfare perspective, bunching only really matters if it is a result of upward manipulation. In Section V, I provide evidence that bunching at the threshold weight increases sentence length, suggesting cases are being rounded up. Here I provide further evidence that bunching is a result of upward drug weight manipulation. I investigate manipulation direction by identifying where excess mass occurs post-FSA, using the pre-FSA distribution as a counterfactual for those groups that experienced bunching decreases. This strategy is described in detail in Section VI. B. and illustrated in Figure A.1.

Figure A.7 presents kernel density graphs showing the pre and post distributions for the groups with bunching decreases. While Figure A.7a is noisy and uninformative, A.7b provides evidence of increased case mass for weights just below the threshold. This implies these racial minority cases in the high cost circuits are no longer manipulated to the bunching point.

To systematically compare effects between racial groups, I employ the difference-in-differences model as described in Equation 3 again. I first consider changes to the distribution mass just below the threshold. Tables A.9 and A.10 present the difference-in-difference results for the model using a binary measure of a case being just below the threshold. This dependant variable is varied across specification, starting with a narrow window of 90-100 percent of the threshold weight (but not inclusive of the threshold weight), then using 80-100 and 70-100 measures. I also run a series of regressions to check the distribution just above the threshold weight. Here, being just above the threshold is defined similarly as the just below specifications.

The regression results suggest manipulation is occurring from below. The percent of individual-circuit-minority cases just below the threshold increases post FSA by an amount

between 5 and 7 percentage points compared to White defendants. This result is consistent in magnitude and significance across specifications. Results for conspiracy circuit White defendant cases are insignificant but all positive. When assessing case mass above the threshold, I find coefficients that are all insignificant and close to zero or negative, implying that manipulation is not being driven by convenience or any other downward rounding behavior.

Interpretation and Mechanisms

Why does the circuit split in conspiracy charging law drive a race gap in drug weight manipulation? To understand the connection between race, manipulation, and prosecutor incentives, I first break down the results for each circuit-race group combination and refer to the model predictions to interpret the findings.

1. **White defendants in individual circuits** experienced essentially no change in bunching. Recall no change in bunching may be indicative of one of two effects. Either most defendants in the group get bunched all the time, likely due to low manipulation costs, or most defendants don't get bunched at all, indicating high manipulation costs. Figure A.4 shows that pre-FSA bunching was low for this group, indicating most defendants are likely not bunched before or after the FSA. Then under the individual method, White defendants must have high costs to manipulation.

2. **Minority defendants in individual circuits** experienced significant decreases in bunching. Decreases in bunching indicate these defendants had low enough costs to be bunched before the FSA, but costs that are too high to justify bunching post-FSA when mandatory minimum benefits decrease, leading to less manipulation after the FSA. Comparing each race group within individual circuits, the FSA response suggests that prosecutors face lower costs in manipulating drug weights for minority defendants.

3. **White defendants in conspiracy circuits** also experienced significant decreases in bunching.

4. **Minority defendants in conspiracy circuits** experienced increases in bunching.

Recall that the model suggests that increases in bunching will only occur if costs to manipulation are low and the optimal sentence to the prosecutor s^* is mid-ranged. This corroborates the idea that costs to manipulation are lower for minority defendants. It also suggests that individual discrimination is likely not the sole driver of effects since discrimination would be reflected with a higher s^*.

Comparing effects within and across circuit types, the results suggest that manipulation costs are lower for minority defendants compared to White counterparts. Why might manipulation costs be lower for Black and Hispanic defendants? Here I posit four plausible mechanisms. The first is the ease of connecting defendants to additional drug weights within a conspiracy may vary by race. That is, it may be easier for a prosecutor to find evidence of additional drugs to connect to the individual if they are a minority compared to if they are White. The second mechanism is similar; it may be easier to find evidence that minorities are involved in a conspiracy at all compared to Whites. Third, pretrial detention has been shown to have significant impacts on case outcomes and to disproportionately affect racial minorities (Leslie & Pope 2018; Dobbie et al. 2018). If racial minorities are experiencing more pretrial detention and doing so reduces their bargaining power, prosecutors may have an easier time manipulating drug weights higher compared to Whites who are less likely to be detained before trial. Finally, if minority defendants have lower quality legal counsel, bargaining to to the bunching point may be less costly for the prosecutor. Regardless of which channel drives the cost difference, the results signify that racial minorities are disadvantaged in sentencing for committing the same types of crimes as White counterparts.

I also argue that these results indicate that the race disparity in drug weight manipulation will exist regardless of whether prosecutors practice individual discrimination. Imagine the disparity in charges was driven entirely by individual prosecutor discrimination, either taste-based, statistical, or some combination. Then the gap between race groups should be proportional between circuit groups both pre and post-FSA. This is clearly not the case, with significant differences in the racial bunching between the two methods of conspiracy charging both pre and post-FSA. Likewise, the differential effects of the FSA on the race gap across circuit groups indicates that racial disparities are at least in part driven by systemic factors.

Robustness

In this section I perform a series of robustness checks to test the strength and validity of the main results. Specifically, I consider whether the findings survive changes to measurement definitions, I increase the sample to include higher and lower drug weight cases, I consider whether results are driven by any one specific drug type, I check if linear trends affect results, I check whether results are driven by outliers, and I consider whether differences among racial groups can be explained by several other potential sources.

Table A.11 shows the results of the main effects of the FSA on bunching using alternative definitions of bunching. I consider a 2-percent window, including all weights at the bunching point and up to 102 percent of the threshold weight, a 5-percent window, and a simple 10 gram window. I find similar results to the main ones, with slightly smaller and less significant effects for individual circuits and larger, more significant effects for conspiracy circuits. Table A.12 considers the main effects for larger samples of cases. I consider a sample with charging weights between 30 percent and 500 percent of the threshold weight. Note that this still excludes the lower threshold, which may be impacted by the FSA. I then consider the full distribution of charging weights. Across these specifications the main results survive, with larger effects in individual circuits and lower, less significant effects in conspiracy circuits. I also consider the manipulation direction analysis for these larger weight windows and find similar results to the main ones.

There may be concern that results are driven by a specific drug type rather than race or circuit type. To alleviate this concern, I graph the main histograms for minority defendants across each circuit type by each of the five drug types in Figures A.8 and A.9. The effects are consistent across each drug type; they show significant decreases in bunching in the individual circuits for every drug type and significant increases in bunching in the conspiracy circuits for each drug type, besides meth which had a null effect. I omit the White graphs which are noisier due to fewer observations, but still largely corroborate the main findings across each drug type.

I also check whether directional manipulation results are driven by the specific weight cutoffs I chose. I consider the main regression showing excess mass below the bunch point

post-FSA in high cost circuits for minority defendants with a series of different dependent variables. I graph the the FSA effect coefficient with 95 confidence intervals for each regression in Figure A.10. Each point on the graph gives an increase in the percent window range of the dependent variable. I find that regardless of the dependent variable definition, the results hold, though very large windows lose significance.

I find the main bunching results and manipulation direction results are robust to a series of time trend controls. Specifically, I consider the models again with a specification including a linear time trend, one including district-specific time trends, and one including race-specific time trends. I find the results keep similar significance and magnitude throughout each of these specifications. To check that results are not driven by gang involvement, I also drop districts with the highest level of Mexican Cartel involvement according to the 2019 DEA National Drug Threat Assessment. This drops a total of seven districts, all of which are in individual method circuits. Results are robust to excluding these high cartel involved districts.[26] In addition, I consider whether results are driven by a handful of outliers or specific districts. Instead, I find results are fairly constant across districts within circuit groups.

Finally, I consider whether the geographic heterogeneity in FSA effects may be explained by other plausible factors outside of the circuit split. I consider differences in political preference as measured by percent voting in the 2016 presidential election, racial animus, general arrest patterns, and state level prison severity. None of these can explain the heterogeneity well.

VII Conclusion

For federal drug trafficking crimes, Black and Hispanic individuals receive harsher charges and higher sentences. In this paper, I provide evidence that racial disparities arise due to prosecutor drug weight manipulation, that minorities are receiving longer sentences due to manipulation, that prosecutors' manipulation decisions are significantly impacted by charg-

[26]Regression tables for these linear trend and district exclusion checks are omitted but are available upon request.

ing law, and that the efficacy of a disparity-reducing policy is dependant upon the charging law framework prosecutors are acting under. I provide evidence that these disparities are not driven by differences in defendant behavior or unobserved case characteristics, and that disparities occur across each of the five major drug types. Together, these results suggest that systemic factors, such as charging law, have significant impacts on racial disparities in drug sentencing.

I utilize the bunching point in drug weight charges at the mandatory minimum threshold weight to identify prosecutor drug weight manipulation. I use a conceptual model of prosecutor choice to make predictions and back out interpretations of the effects of the FSA as a policy intervention. I then compare bunching differences across race and circuit type using US Attorney spell propensity to bunch as a proxy for bunching probability. I next perform a second analysis investigating the effects of the First Step Act on prosecutor charging decisions. The FSA effectively reduced the expected benefit of bunching for prosecutors by decreasing the expected impact of a mandatory minimum sentence. I then compare the effects of this policy across race and across federal circuits with differences in costs to manipulation.

The first analysis shows that prosecutors disproportionately manipulate drug weights for racial minorities compared to White counterparts. I then provide evidence that manipulation has significant impacts to sentence length. The second analysis shows that decreasing the impact of mandatory minimums significantly reduces the racial disparity in bunching when costs to drug weight manipulation are high. But when costs to manipulation are low, the FSA actually increases the disparity. This highlights the importance of legal framework in crafting effective policy.

Taken together, the results give evidence that racial minorities are significantly disadvantaged in the sentencing process compared to White defendants with similar crimes, criminal history levels, and acting at similar times and in similar areas. Furthermore, the results imply that disadvantages are directly linked to prosecutor incentives, which are strongly impacted by charging law. Under the framework of the model, the heterogeneous FSA effects are consistent with the idea that prosecutors consider the costs to manipulation lower for Black and Hispanics than for Whites. This is why bunching increases for minorities in the

conspiracy circuits following the FSA; the cost of manipulation is still lower than the benefit of a higher sentence for the prosecutor. In individual circuits, the costs to manipulation are higher overall, leading to decreases in minority bunching. This means that to lower mandatory minimum sentences for minorities and reduce the race gap in charges and sentences, the costs to manipulation should be raised by implementing the individual method of charging in each circuit.

While I do not formally measure discrimination in this paper, it is worth considering whether these results are indicative of discriminatory outcomes. My quantitative design, inclusion of controls, and descriptive statistics largely rule out criminal behavior as the source of disparity. This implies the race gap in charges is driven either by individual discrimination or systemic factors. I am unable to measure any sort of individual discrimination as I do not observe individual prosecutors. However, I show that disparities are at least in part driven by systemic factors. The system disadvantages minorities based on the ability of the prosecutor to gather evidence tying an individual to additional drug quantities, the ability of the prosecutor to connect an individual to a conspiracy, having lower quality legal counsel, or a combination of these effects. This may thought of as technological systemic discrimination as defined in Bohren et al. (2022).

This paper raises several important and exciting questions for future research. Particularly, it highlights the importance of understanding prosecutor incentives and the way in which institutional framework can shape those incentives. More work is needed to understand how prosecutors consider costs and benefits by defendant type. Individual prosecutor data will be key to uncovering these effects, decomposing individual and systemic discrimination, and finding effective policy solutions to unwarranted racial disparities in the criminal justice system.

References

Aizer, A., & Doyle, J. J., Jr. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *Quarterly Journal of Economics*, *139*(2), 759-803.

Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*, *133*(4), 1885-1932.

Arora, A. (2019). *Too tough on crime? the impact of prosecutor politics on incarceration* (Working Paper). American Economic Association.

Bebchuk, L. A. (1984). Litigation and settlement under imperfect information. *RAND Journal of Economics*, *15*(3), 404-415.

Bhuller, M., Dahl, G. B., Loken, K. V., & Magne, M. (2016). *Incarceration, recidivism and employment* (Working paper). National Bureau of Economic Research Working Paper 22648.

Bjerk, D. (2005). Making the crime fit the penalty: The role of prosecutor discretion under mandatory minimum sentencing. *Journal of Law and Economics*, *48*(2), 591-625.

Bjerk, D. (2017a). Mandatory minimum policy reform and the sentencing of crack cocaine defendants: An analysis of the fair sentencing act. *Journal of Empirical Legal Studies*, *14*(2), 370-396.

Bjerk, D. (2017b). Mandatory minimums and the sentencing of federal drug crimes. *Journal of Legal Studies*, *46*(1), 93-128.

Bohren, J. A., Hull, P., & Imas, A. (2022). *Systemic discrimination: Theory and measurement* (Working Paper).

Boylan, R. T. (2004). Salaries, turnover, and performance in the federal us justice system. *Journal of Law and Economics*, *47*(1), 75-92.

Boylan, R. T. (2005). What do prosecutors maximize? evidence from the careers of us attorneys. *American Law and Economics Review*, *7*(2), 379-402.

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, *134*(3), 1405-1454.

Dahl, G. B., Kostol, A. R., & Mogstad, M. (2014). Family welfare cultures. *Quarterly Journal of Economics*, *129*(4), 1711-52.

DEA. (2019). *National drug threat assessment.* `https://www.dea.gov/sites/default/files/2020-01/2019-NDTA-final-01-14-2020_Low_Web-DIR-007-20_2019.pdf`.

Didwania, S. (2020). *Mandatory minimums and federal sentences* (Working Paper). Temple University.

Didwania, S. H. (2020). Mandatory minimum entrenchment and the controlled substances act. *Ohio State Journal of Criminal Law*, *18*.

Di Tella, R., & Schargrodsky, E. (2013). Criminal recidivism after prison and electronic monitoring. *Journal of Political Economy*, *121*(1), 28-73.

Dobbie, W., Goldin, J., & Yang, C. S. (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, *108*(2), 201-40.

Fischman, J. B., & Schanzenbach, M. M. (2012). Racial disparities under the federal sentencing guidelines: The role of judicial discretion and mandatory minimums. *Journal of Empirical Legal Studies*, *9*(4), 729-764.

Frandsen, B. R. (2017). Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. In *Regression discontinuity designs: Theory and applications.* Bingly, UK: Emerald Publishing Limited.

Glaeser, E. L., Kessler, D. P., & Piehl, A. M. (2000). What do prosecutors maximize? an analysis of federalization of drug crimes. *American Law and Economics Review*, *2*(2), 259-290.

Goncalves, F., & Mello, S. (2021). A few bad apples? racial bias in policing. *American Economic Review*, *111*(5), 1406-41.

Gordon, S. C., & Huber, G. A. (2002). Citizen oversight and the electoral incentives of criminal prosecutors. *American Journal of Political Science*, *46*(2), 334-351.

Justice manual (Tech. Rep.). (2018). Washington, DC: United States Department of Justice.

Kleven, H. (2016). Bunching. *Annual Review of Economics*, *8*, 435-464.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *American Economic Review*, *96*(3), 863-76.

Kurth, E. (2020). Drug conspiracy sentencing and social injustice. *The University of Colorado Law Review*, *91*(4), 1216-1246.

Landes, W. (1971). An economic analysis of the courts. *Journal of Law and Economics*, *14*(1), 61-107.

Leslie, E., & Pope, N. (2018). The unintended impact of pretrial detention on case outcomes: Evidence from nyc arraignments. *Journal of Law and Economics*(60), 529-557.

Lynch, M. (2016). *Hard bargains: The coercive power of drug laws in federal court.* New York: Russell Sage Foundation.

McKinley, E. (2019). The importance of drug quantity in federal sentencing: How circuit courts should determine the mandatory minimum sentence for conspiracy to distribute controlled substances in light of united states v. stoddard. *The University of Cincinnati Law Review*, *87*(4), 1145-1165.

Mueller-Smith, M. (2015). *The criminal and labor market impacts of incarceration* (Working paper).

Rasmusen, E., Raghav, M., & Ramseyer, M. (2009). Convictions versus conviction rates: The prosecutor's choice. *American Law and Economics Review*, *11*(1), 47-78.

Rehavi, M. M., & Starr, S. B. (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy*, *122*(6), 41320-1354.

Rhodes, W., Kling, R., Luallen, J., & Dyous, C. (2015). *Federal sentencing disparity: 2005-2012* (Working paper). BJS Working Paper Series.

Robertson, C., Baughman, S. B., & Wright, M. S. (2019). Race and class: A randomized experiment with prosecutors. *Journal of Empirical Legal Studies*, *16*(4), 807–847.

Silveira, B. S. (2017). Bargaining with asymmetric information: An empirical study of plea negotiations. *Econometrica*, *85*(2), 419-452.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using google search data. *Journal of Public Economics*, *118*, 26-40.

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

United States v. Banuelos, 332 F.3d 700. (9th Cir. 2003). Retrieved from `https://casetext.com/case/us-v-banuelos-2`

United States v. Ellis, 868 F.3d 1155. (10th Cir. 2017). Retrieved from `https://casetext.com/case/united-states-v-ellis-138`

United States v. Gibson, 874 F.3d 544. (6th Cir. 2017). Retrieved from `https://casetext.com/case/united-states-v-gibson-226`

United States v. Haines, 803 F.3d 713. (5th Cir. 2015). Retrieved from `https://casetext.com/case/united-states-v-haines-20`

United States v. Knight, 342 F.3d 697. (7th Cir. 2003). Retrieved from `https://casetext.com/case/us-v-knight-42`

United States v. Phillips, 349 F.3d 138. (3rd Cir. 2003). Retrieved from `https://casetext.com/case/us-v-phillips-70`

United States v. Pizarro, 772 F.3d 284. (1st Cir. 2014). Retrieved from `https://casetext.com/case/united-states-v-pizarro-10`

United States v. Rangel, 697 F.3d 795. (4th Cir. 2015). Retrieved from `https://casetext.com/case/united-states-v-rangel-59`

United States v. Robinson, 547 F.3d 632. (6th Cir. 2008). Retrieved from `https://casetext.com/case/us-v-robinson-475`

United States v. Stiger, 413 F.3d 1185. (10th Cir. 2005). Retrieved from `https://cite.case.law/f3d/413/1185/`

United States v. Stoddard, 892 F.3d 1203. (D.C. Cir. 2018). Retrieved from https://casetext.com/case/united-states-v-stoddard-11

United States v. Young, 847 F.3d 328. (6th Cir. 2017). Retrieved from https://casetext.com/case/united-states-v-young-585

Tuttle, C. (2021). *Racial disparities in federal sentencing: Evidence from drug mandatory minimums* (Working Paper). University of Texas.

USSC. (2019). *First step act.* https://www.ussc.gov/sites/default/files/pdf/training/newsletters/2019-special_FIRST-STEP-Act.pdf.

USSC. (2021). *Drug primer.* https://www.ussc.gov/sites/default/files/pdf/training/primers/2021_Primer_Drugs.pdf.

Yang, C. (2016). Resource constraints and the criminal justice system: Evidence from judicial vacancies. *American Economic Journal: Economic Policy*, *8*(4), 289-332.

A Appendix

Notes on Enhancement and Safety Valve Provision

Sentence enhancement refers to an increase in a punishment given when the crime involves a specified undesirable quality or behavior. In context of the First Step Act (FSA), the enhancement refers to an increased mandatory minimum length given for offenders with past drug or violent crimes. Prior to FSA, offenders with a prior drug or violent felony conviction were eligible for an enhancement that would double the mandatory minimum sentence length. For enhancement to be applied, the prosecution must explicitly file for a notice of enhancement under 21 U.S.C. §851 and these must be confirmed by a jury.[27] This means that if an offender is convicted of trafficking drugs at a weight above the high threshold and they receive the enhancement, they would be subject to a mandatory minimum of 20 years rather than the typical statutory 10 years. FSA changed enhancement procedure in two ways. First, it decreased the high threshold penalty from 20 years to 15 years. Second, the qualification for enhancement changed from a prior drug or violent felony conviction to a "serious" prior drug or violent felony conviction. Here, "serious" is defined as a drug or violent crime for which the offender served a term of imprisonment for longer than 12 months.[28] Taken together, these two effects imply there are less individuals who qualify for enhancement and the penalty for enhancement at the higher threshold is decreased.

The second major sentencing revision introduced by the FSA is an expansion of the safety valve provision. Safety valve provides a means whereby offenders can receive a sentence below the mandatory minimum even if the quantity of drugs they are charged with exceeds the threshold. The intent of the provision was to provide relief for first-time, non-violent offenders. To qualify for the safety valve provision, the defense must provide proof at a preponderance of evidence level that the defendant did not use violence or possess a dangerous weapon in connection with the crime, the offense did not result in death or serious injury

[27] Prior to *Alleyne v. United States* enhancement factors were to be determined by the judge. However, the observation period only considers cases post *Alleyne*, so all enhancements in this study come by approval of a jury.

[28] The details of which crimes qualify as "serious" are more technical than explained here, but include most drug or violent crimes for which the offenders was imprisoned for over 12 months. See Public Law 115-391 for more details of seriousness as well as which statutes qualify as a violent crime.

of any persons, the defendant did not act as a leader or organizer, the defendant cooperates with law enforcement and gives all known information of the crime (if applicable), and the defendant does not have more than one criminal history point. One criminal history point translates to having one or less prior offense, and this offense must have resulted in a sentence of less than 60 days incarceration. This criminal history stipulation is especially binding as many non-violent offenders have more than one criminal history point.[29] FSA expanded the safety valve provision to include individuals with higher criminal history points (up to four, excluding one point offenses). However, offenders with a prior conviction that resulted in more than 12 months incarceration or a violent prior conviction that resulted in more than 60 days incarceration are still exempt.

[29]In 2017, around 66% of all federal trafficking cases involved offenders with criminal history points over one point.

Tables

Table A.1: Mandatory Minimum Drug Weight Thresholds

Drug Type	5-Year Threshold Amount (in grams)	10-Year Threshold Amount (in grams)
Cocaine	500	5,000
Crack	28	280
Heroin	100	1,000
Marijuana	100,000	1,000,000
Meth	5	50

Table A.2: US Attorney Data - Summary Statistics

	mean	sd
number of US attorneys per district	3.168	(0.905)
months served	48.28	(26.40)
number of cases	134.2	(150.9)
bunching propensity measure	9.750	(12.84)
max bunching propensity within district	16.41	(15.05)
min bunching propensity within district	2.752	(2.822)
# of Attorneys		282

Table A.3: USSC Drug Case Data - Summary Statistics

	All Cases	Individual Circuits	Conspiracy Circuits
Panel A. Drug Types			
cocaine	0.257	0.259	0.249
	(0.437)	(0.438)	(0.433)
crack	0.109	0.108	0.113
	(0.312)	(0.311)	(0.317)
heroin	0.185	0.167	0.242
	(0.389)	(0.373)	(0.428)
marijuana	0.101	0.117	0.0516
	(0.301)	(0.321)	(0.221)
meth	0.348	0.349	0.344
	(0.476)	(0.477)	(0.475)
Panel B. Defendant Characteristics			
white	0.245	0.232	0.286
	(0.430)	(0.422)	(0.452)
black	0.296	0.282	0.340
	(0.457)	(0.450)	(0.474)
Hispanic	0.458	0.485	0.374
	(0.498)	(0.500)	(0.484)
female	0.152	0.155	0.141
	(0.359)	(0.362)	(0.348)
age	36.15	36.02	36.56
	(9.764)	(9.831)	(9.541)
criminal history points	4.307	4.221	4.575
	(5.406)	(5.450)	(5.256)
number of charges	2.384	2.428	2.244
	(1.171)	(1.160)	(1.193)

college	0.207	0.209	0.201
	(0.405)	(0.407)	(0.401)
illegal alien	0.141	0.144	0.135
	(0.349)	(0.351)	(0.342)
Panel C. Outcomes			
percentage weight	0.920	0.922	0.914
	(0.274)	(0.275)	(0.274)
conspiracy charge	0.503	0.487	0.554
	(0.500)	(0.500)	(0.497)
baseline charge level	28.31	28.35	28.18
	(2.940)	(2.964)	(2.860)
10 year MM sentence imposed	0.212	0.211	0.216
	(0.409)	(0.408)	(0.412)
safety valve applied	0.298	0.313	0.252
	(0.457)	(0.464)	(0.434)
assistance to government	0.353	0.346	0.377
	(0.478)	(0.476)	(0.485)
sentence length (in months)	75.32	74.79	76.96
	(60.37)	(60.26)	(60.69)
Share	1	0.759	0.241
Observations	12,722	9,643	3,079

Statistics are means with standard deviations in parentheses

Table A.4: Bunching Propensity Randomization Check

	(1) bunch	(2) bunch_score	(3) bunch_score	(4) bunch_score
F-Value:	4.88	0.99	1.03	1.48
F-Test:	0.000	0.475	0.438	0.155
mean:	9.586	9.461	9.430	8.508
Population Controls	X	X	X	X
Fixed Effects	X	X	X	X
Attorney # of Cases	All	All	≥ 25	≥ 50
R Squared	0.184	0.871	0.905	0.935
N	12,216	12,216	10,927	8,553

Notes: Here I regress the residualized bunching propensity measure on defendant characteristics. Each specification includes year, month, and district fixed effects and measures of racial population percentages by year. Covariates included in the regression are drug type, sex criminal history points, age and age squared, a binary measure for college, and a binary measure for illegal alien.

Table A.5: Racial Disparity in Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)
Black	14.72***	14.60***	15.72***	15.27***	15.15***	14.59***
	(2.651)	(2.869)	(3.641)	(2.449)	(2.623)	(3.314)
Hispanic	4.667***	5.360***	6.697***	6.085***	6.811***	7.636***
	(1.733)	(1.923)	(2.230)	(1.719)	(1.922)	(2.268)
bunch_score	-0.0895	-0.134	-0.185	0.00244	-0.0677	-0.233
	(0.129)	(0.110)	(0.197)	(0.130)	(0.127)	(0.210)
Black*bunch_score	0.385***	0.324**	0.247*	0.0663	-0.0109	0.211
	(0.140)	(0.144)	(0.126)	(0.163)	(0.174)	(0.243)
Hispanic*bunch_score	0.287***	0.221**	0.294**	-0.0564	-0.122	0.00584
	(0.0832)	(0.0943)	(0.117)	(0.111)	(0.123)	(0.177)
Includes Bunching Point	X	X	X			
Attorney # of Cases	All	≥ 25	≥ 50	All	≥ 25	≥ 50
R Squared	0.310	0.315	0.316	0.348	0.354	0.355
N	12,216	10,927	8,553	11,045	9,889	7,825

Notes: The dependant variable is sentence length at disposition. All specifications include the full set of controls; defendant characteristics, time and district fixed effects, and race population controls. The first three specifications include the bunching point while columns 4-6 omit all cases at the bunching point. This tests whether disparities in sentence length between races is driven by drug weight manipulation.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Changes in Bunching - Black and Hispanic Separate

	Individual Circuits		**Conspiracy Circuits**	
	(1)	(2)	(3)	(4)
Black	0.0412*	0.0244	-0.0415*	-0.0385*
	(0.0214)	(0.0169)	(0.0240)	(0.0224)
Hispanic	0.00897	0.00629	-0.0772***	-0.0567**
	(0.0187)	(0.0101)	(0.0261)	(0.0255)
fsa	0.00514	0.0107	-0.0659***	-0.0863**
	(0.0144)	(0.0157)	(0.0225)	(0.0313)
Black*fsa	-0.0513***	-0.0534***	0.0960***	0.109***
	(0.0191)	(0.0184)	(0.0242)	(0.0313)
Hispanic*fsa	-0.0301**	-0.0533***	0.0956***	0.0671
	(0.0150)	(0.0158)	(0.0337)	(0.0448)
Covariates	X	X	X	X
Circuit Decision Controls	X	X	X	X
Population Controls		X		X
Fixed Effects		X		X
R Squared	0.078	0.239	0.062	0.167
N	9,643	9,637	3,079	3,079

Standard errors in parentheses

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

Notes: The dependant variable is a binary variable that equals one if the charging weight is exactly equal to the mandatory minimum threshold weight.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Robustness Check - Omitting High and Low Case-per-month Attorneys

	(1) bunch	(2) sent length	(3) bunch	(4) sent length
Black	-0.0244*	14.82***	-0.0194	14.48***
	(0.0123)	(2.857)	(0.0124)	(2.569)
Hispanic	-0.0194**	4.270**	-0.0171*	4.525**
	(0.00961)	(1.900)	(0.00932)	(1.895)
bunch_score	-0.00119	-0.292*	0.000389	-0.172
	(0.00146)	(0.148)	(0.00147)	(0.115)
Black*bunch_score	0.00369***	0.303**	0.00296**	0.276**
	(0.00116)	(0.138)	(0.00123)	(0.120)
Hispanic*bunch_score	0.00156**	0.294***	0.00115	0.213**
	(0.000769)	(0.0942)	(0.000767)	(0.0857)
Attorney Cases per Month	≥ 0.5	≥ 0.5	≤ 9	≤ 9
R Squared	0.191	0.316	0.192	0.311
N	11,190	11,190	10,701	10,701

Notes: All specifications include the full set of controls; defendant characteristics, time and district fixed effects, and race population controls. The first two specifications omit US Attorney spell cases with the bottom 10 percent of cases per month. Columns 3 and 4 omit the top 10 percent.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Robustness Check - Bunching Disparity using a Binary Bunching Score

	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.00784	0.0107	0.00755	-0.00484	-0.00321	-0.0150
	(0.00922)	(0.0102)	(0.00894)	(0.0111)	(0.0121)	(0.0123)
Hispanic	-0.00419	-0.00365	-0.00729	-0.0269***	-0.0235***	-0.0286***
	(0.00521)	(0.00544)	(0.00601)	(0.00861)	(0.00864)	(0.0103)
bunching binary	0.0169	0.0245*	0.0274	0.0214*	0.0269*	0.00711
	(0.0122)	(0.0140)	(0.0199)	(0.0125)	(0.0152)	(0.0300)
Black*bunching binary	0.0275*	0.0236	0.0291*	0.0201	0.0190	0.0356*
	(0.0147)	(0.0163)	(0.0171)	(0.0154)	(0.0174)	(0.0196)
Hispanic*bunching binary	0.0274**	0.0241**	0.0248*	0.0177	0.0111	0.0181
	(0.0111)	(0.0115)	(0.0143)	(0.0112)	(0.0118)	(0.0140)
Covariates				X	X	X
Population Controls	X	X	X	X	X	X
Fixed Effects	X	X	X	X	X	X
Attorney # of Cases	All	≥ 25	≥ 50	All	≥ 25	≥ 50
R Squared	0.167	0.173	0.169	0.185	0.191	0.184
N	12,216	10,927	8,553	12,216	10,927	8,553

Notes: The dependent variable is the binary measure for being bunched. The key difference from Table 1 is the bunching propensity score measure is now a binary variable rather than a continuous one.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: Cases Just Below the Threshold - **Individual Circuits**

	90-100	80-100	70-100	90-100	80-100	70-100
minority	-0.0244**	-0.0403***	-0.0429***	-0.0272**	-0.0392***	-0.0373**
	(0.00990)	(0.0131)	(0.0152)	(0.0106)	(0.0134)	(0.0159)
fsa	-0.0250*	-0.0245	-0.0275	-0.0334*	-0.0405	-0.0341
	(0.0128)	(0.0160)	(0.0171)	(0.0190)	(0.0253)	(0.0239)
minority*fsa	0.0506***	0.0631**	0.0671**	0.0538***	0.0713***	0.0694***
	(0.0182)	(0.0241)	(0.0259)	(0.0185)	(0.0242)	(0.0260)
pre-fsa mean	0.0964	0.1840	0.2555	0.0964	0.1840	0.2555
Covariates	X	X	X	X	X	X
Circuit Decision Controls	X	X	X	X	X	X
Population Controls				X	X	X
Fixed Effects				X	X	X
R Squared	0.018	0.012	0.011	0.049	0.040	0.040
N	9,643	9,643	9,643	9,637	9,637	9,637

Notes: The dependant variable here is a binary measure of whether a case has a charging weight within the percentage window specified, while still below the threshold weight. Standard errors are clustered at the district level.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: Cases Just Below the Threshold - **Conspiracy Circuits**

	90-100	80-100	70-100	90-100	80-100	70-100
minority	0.0086	0.0165	0.0248	0.0150	0.0286	0.0424*
	(0.0157)	(0.0187)	(0.0232)	(0.0163)	(0.0171)	(0.0209)
fsa	0.0000	0.0146	0.0099	0.0176	0.0514	0.0453
	(0.0233)	(0.0287)	(0.0326)	(0.0256)	(0.0340)	(0.0478)
minority*fsa	-0.0041	-0.0393	-0.0263	-0.0265	-0.0550	-0.0291
	(0.0315)	(0.0391)	(0.0412)	(0.0337)	(0.0459)	(0.0543)
pre-fsa mean	0.0935	0.1923	0.2647	0.0935	0.1923	0.2647
Covariates	X	X	X	X	X	X
Circuit Decision Controls	X	X	X	X	X	X
Population Controls				X	X	X
Fixed Effects				X	X	X
R Squared	0.016	0.006	0.007	0.057	0.048	0.053
N	3,079	3,079	3,079	3,079	3,079	3,079

Notes: The dependant variable here is a binary measure of whether a case has a charging weight within the percentage window specified, while still below the threshold weight. Standard errors are clustered at the district level.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.11: Alternative Bunching Definition Check

	Individual Circuits			Conspiracy Circuits		
	2% range	5% range	10g range	2% range	5% range	10g range
minority	0.00581	0.0131	0.0592***	-0.0600***	-0.0498*	-0.0313
	(0.0120)	(0.0140)	(0.0166)	(0.0203)	(0.0249)	(0.0277)
fsa	-0.0104	-0.0155	0.00928	-0.0862***	-0.105***	-0.163***
	(0.0170)	(0.0200)	(0.0281)	(0.0288)	(0.0308)	(0.0308)
minority*fsa	-0.0349*	-0.0295	-0.0578**	0.114***	0.131***	0.156***
	(0.0188)	(0.0216)	(0.0261)	(0.0332)	(0.0335)	(0.0394)
R Squared	0.203	0.171	0.170	0.147	0.136	0.117
N	9,637	9,637	9,637	3,079	3,079	3,079

Notes: This table replicates the main results varying the definition of bunching. All specifications contain the full set of controls.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.12: Full Weight Regressions

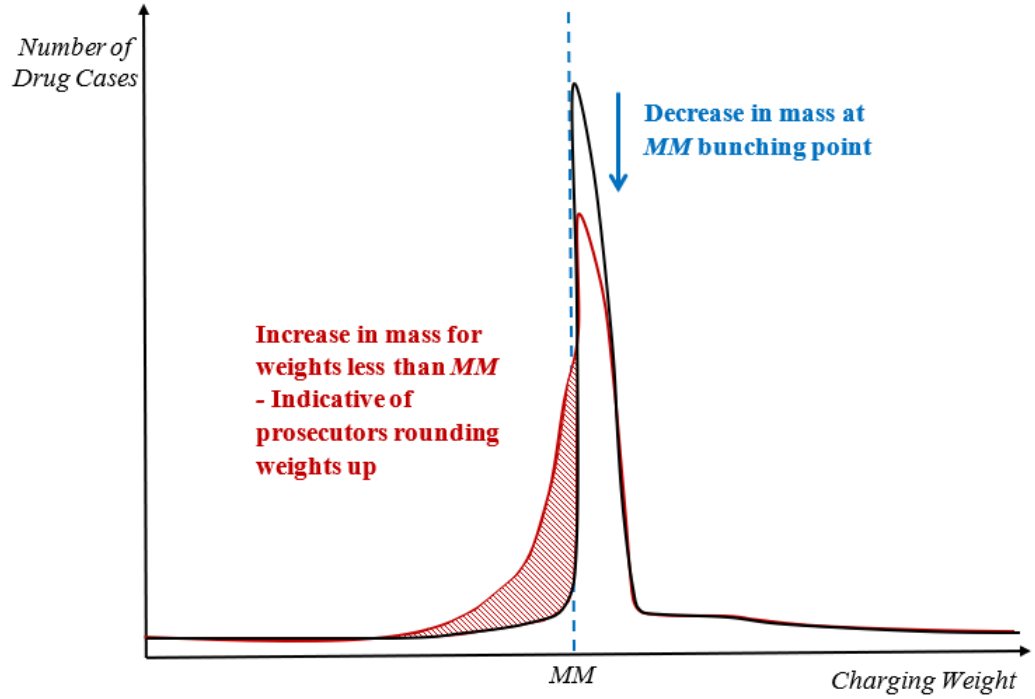
	Individual Circuits		Conspiracy Circuits	
	(1)	(2)	(3)	(4)
minority	0.00907*	0.00218*	-0.0213*	-0.00504*
	(0.00460)	(0.00118)	(0.0113)	(0.00260)
fsa	0.00600	0.00238	-0.0258*	-0.00645*
	(0.00641)	(0.00185)	(0.0132)	(0.00321)
minority*fsa	-0.0271***	-0.00764***	0.0296*	0.00777*
	(0.00683)	(0.00217)	(0.0166)	(0.00402)
Drug Weight Window	30-500	Full	30-500	Full
R Squared	0.134	0.027	0.088	0.017
N	21,980	90,483	6,614	26,949

Notes: This table replicates the main results allowing for a larger window of weights. Columns 1 and 3 give all cases between 30 percent and 500 percent of the threshold weight. Columns 2 and 4 contain the full distribution of cases. All specifications contain the full set of controls.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figures

Figure A.1: Impact of FSA on Charging Weight Distribution

(a) Evidence of Rounding Up



(b) Evidence of Rounding Down

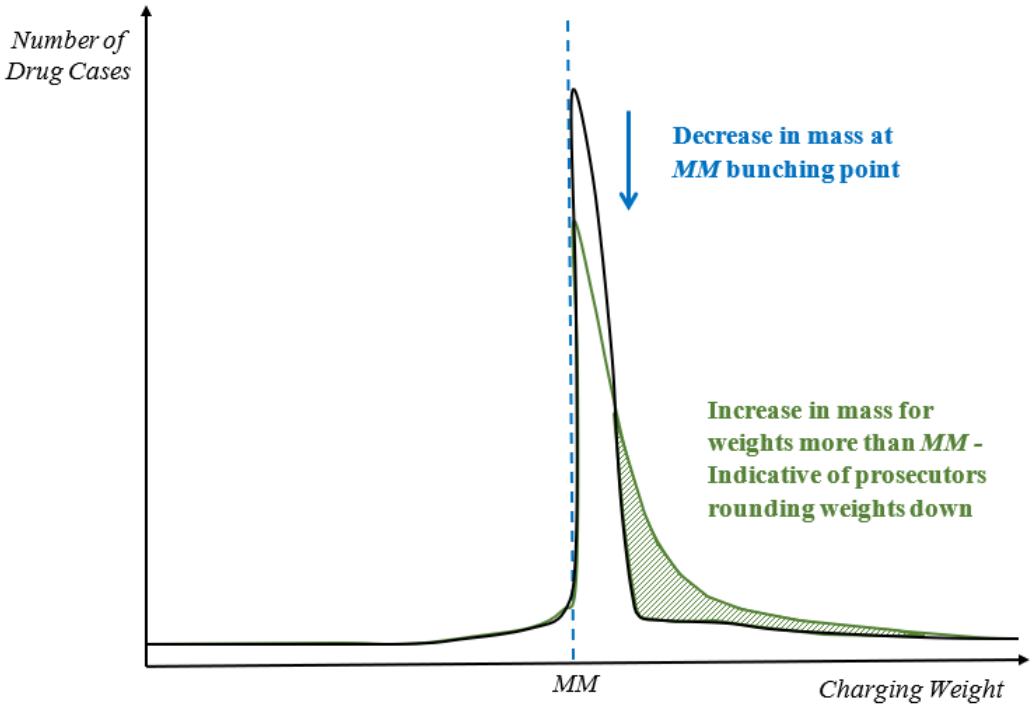


Figure A.2: Percent of Cases Bunched by District

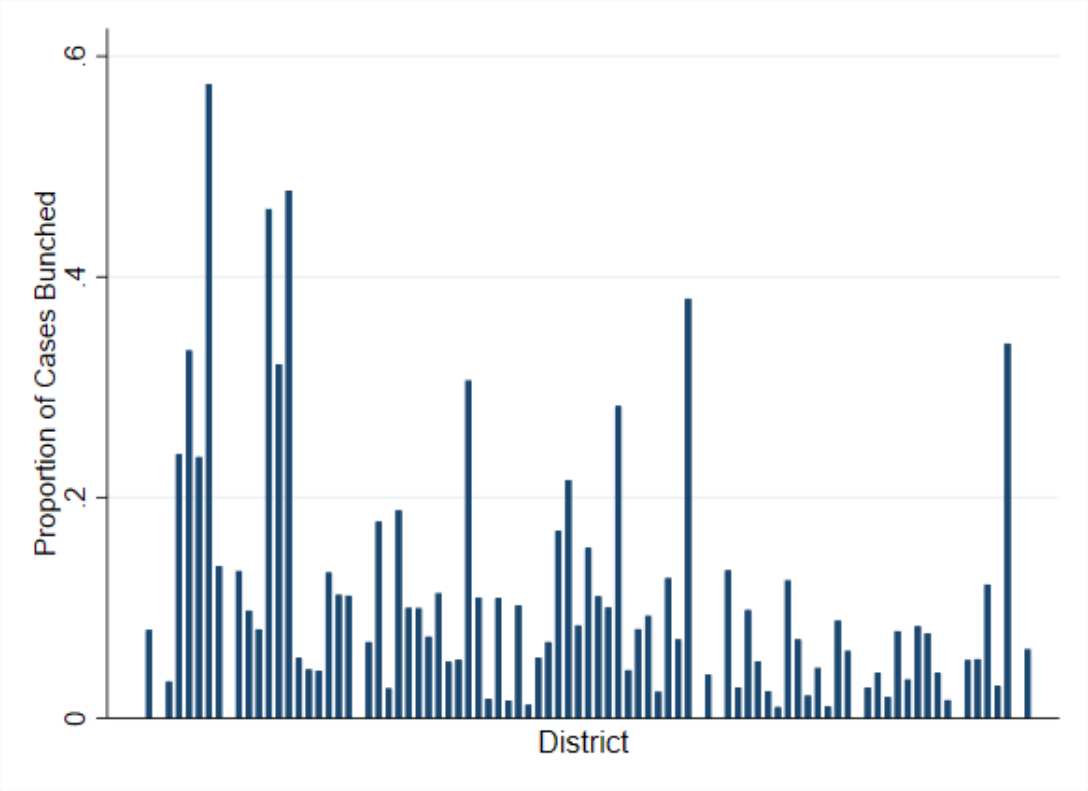


Figure presents a the ratio of cases bunched over the full pre-period across each district. Each bar gives an individual district's average bunching.

Figure A.3: Difference in Max and Min Bunching Propensity by District

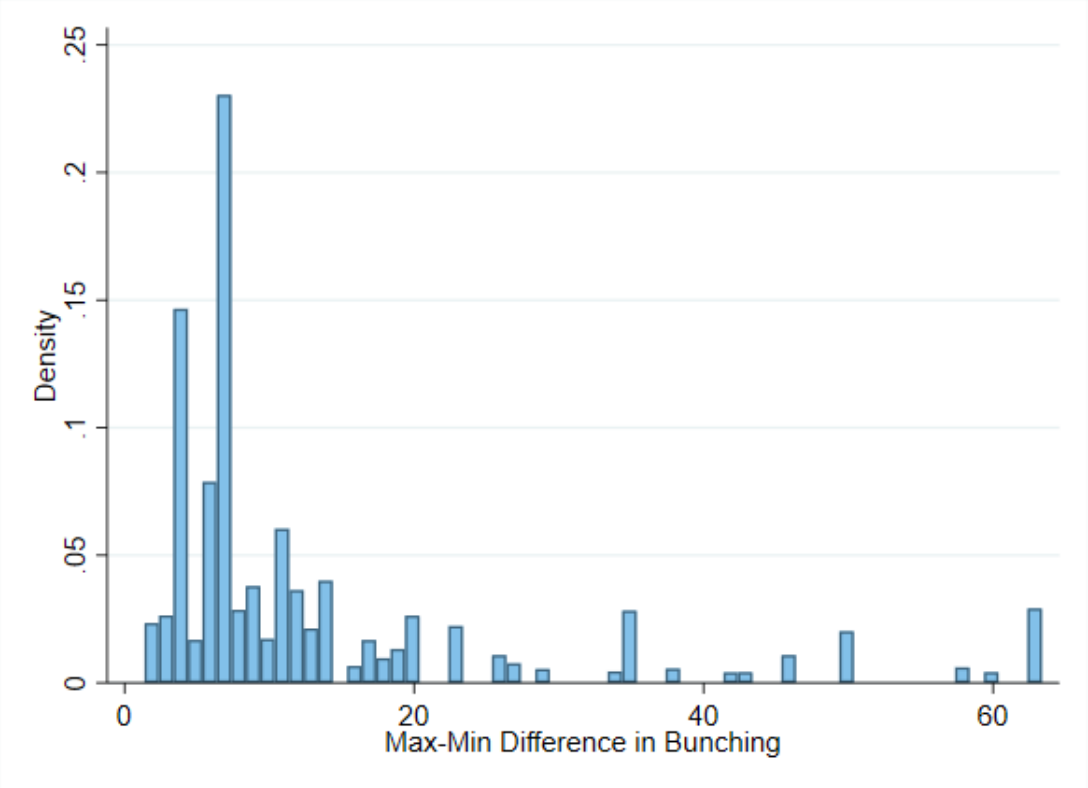
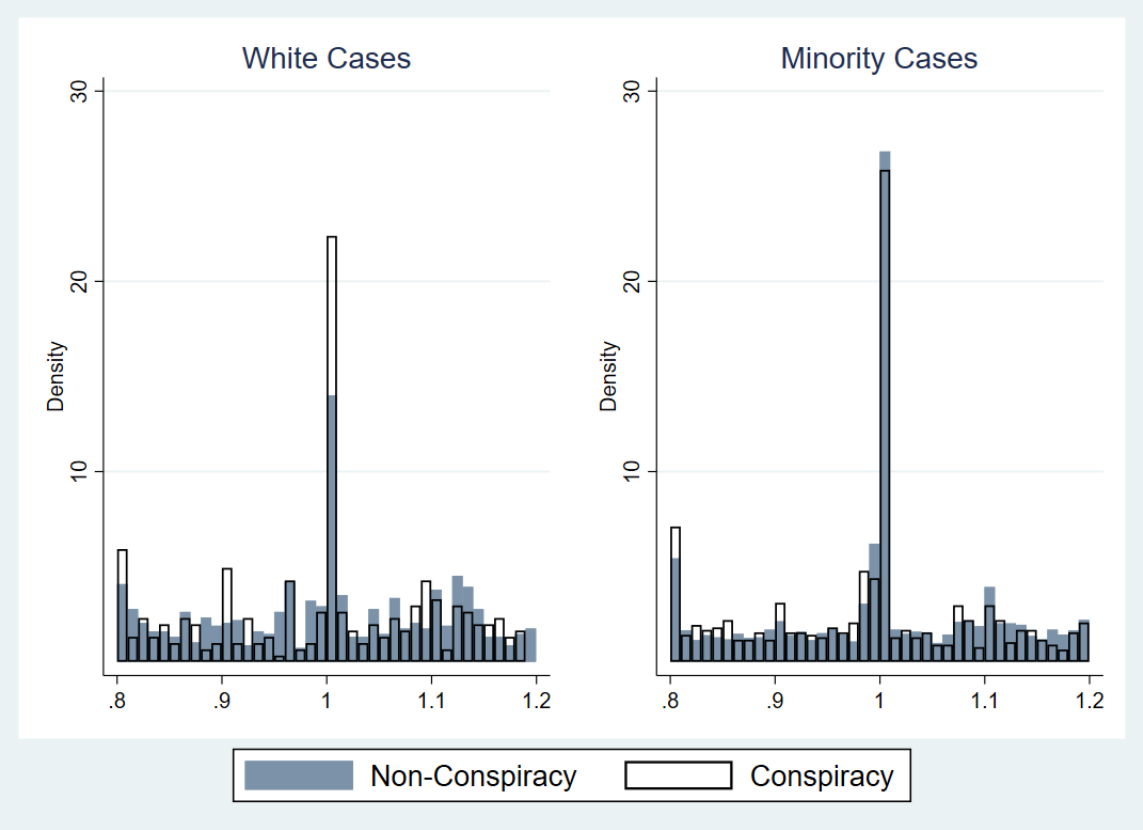


Figure presents a histogram of the difference between the maximum and minimum amounts of bunching for an attorney within each district. The distribution only includes districts with measures 20 or lower. There exist a few outlier districts with much higher max-min differences, creating a long right hand tail in the distribution. For readability, these have been omitted from this figure.

Figure A.4: Pre-FSA Bunching Patterns



These histograms show the bunching patterns for each race-circuit group. This illustrates that pre-FSA, the circuit split appeared to only affect White defendants. This also implies the race gap was large in non-conspiracy circuits pre-FSA while the race gap was much smaller for conspiracy circuits.

Figure A.5: Charged Drug Weight Pre vs Post FSA

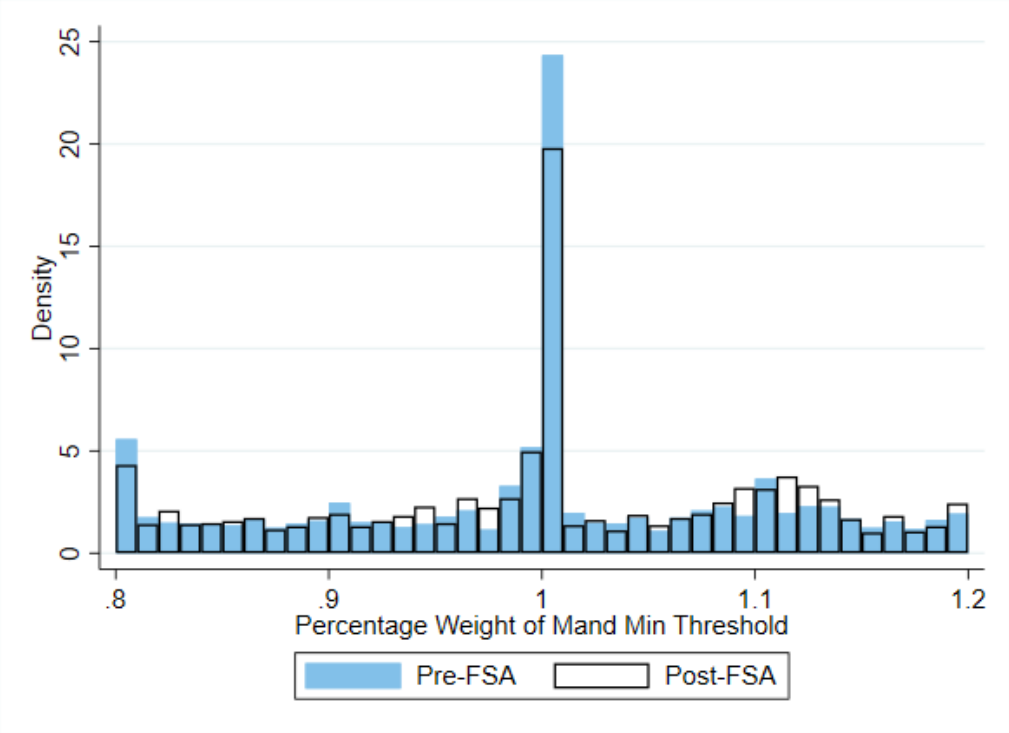
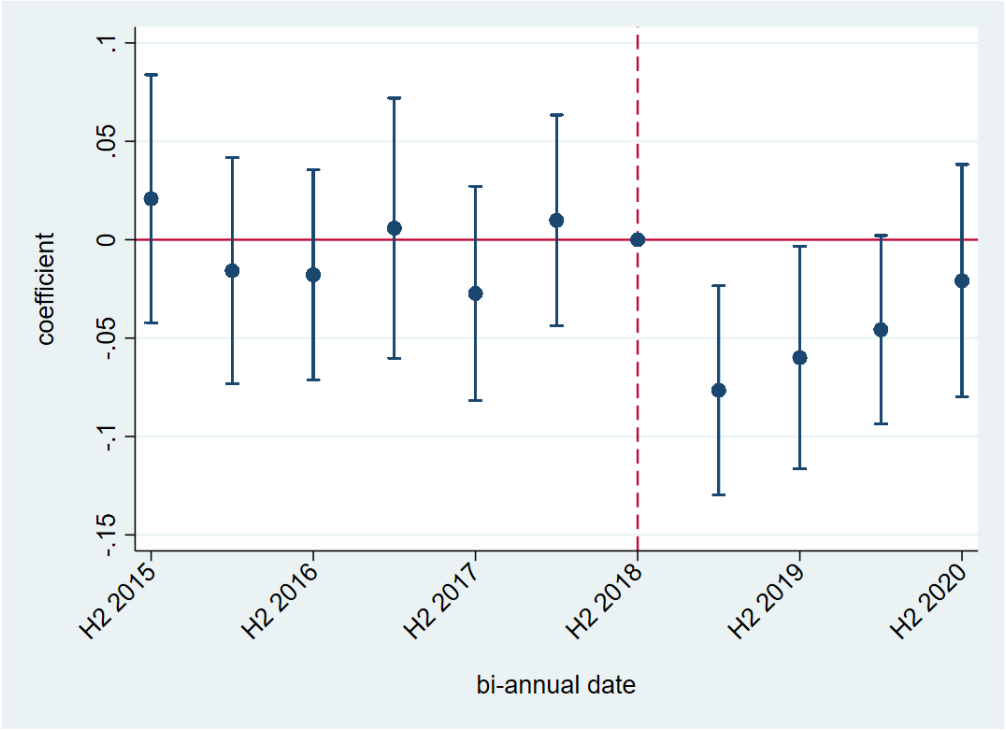


Figure presents a histogram of charge weights for all cases in the sample before the FSA was enacted versus all cases that occurred after.

Figure A.6: Parallel Trends Check

(a) Individual Circuits



(b) Conspiracy Circuits

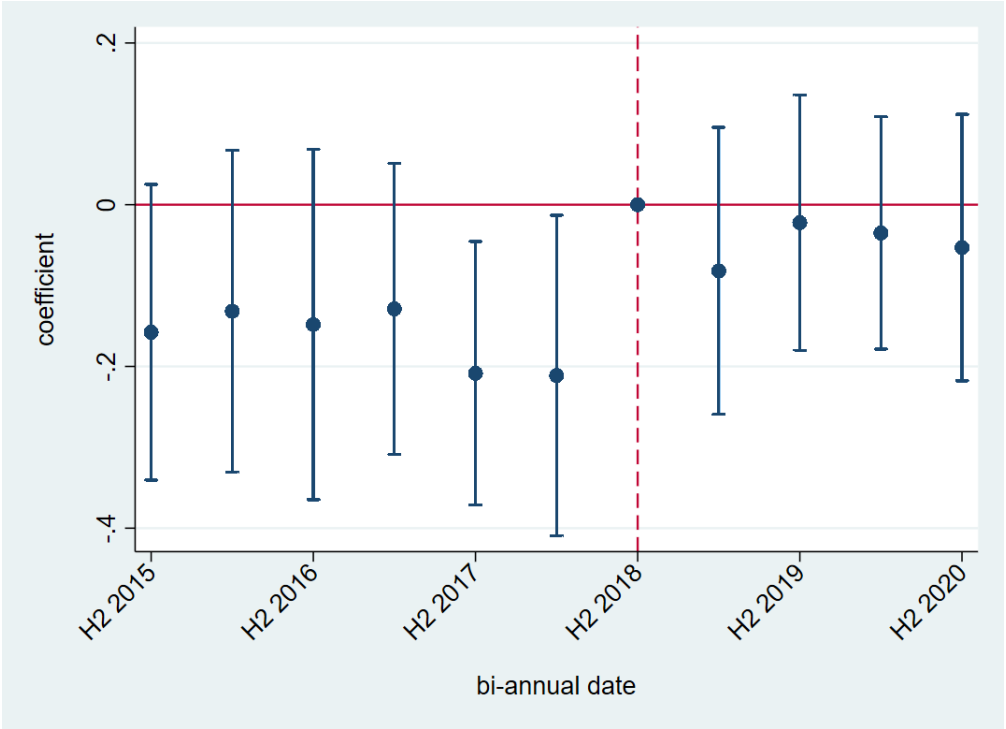
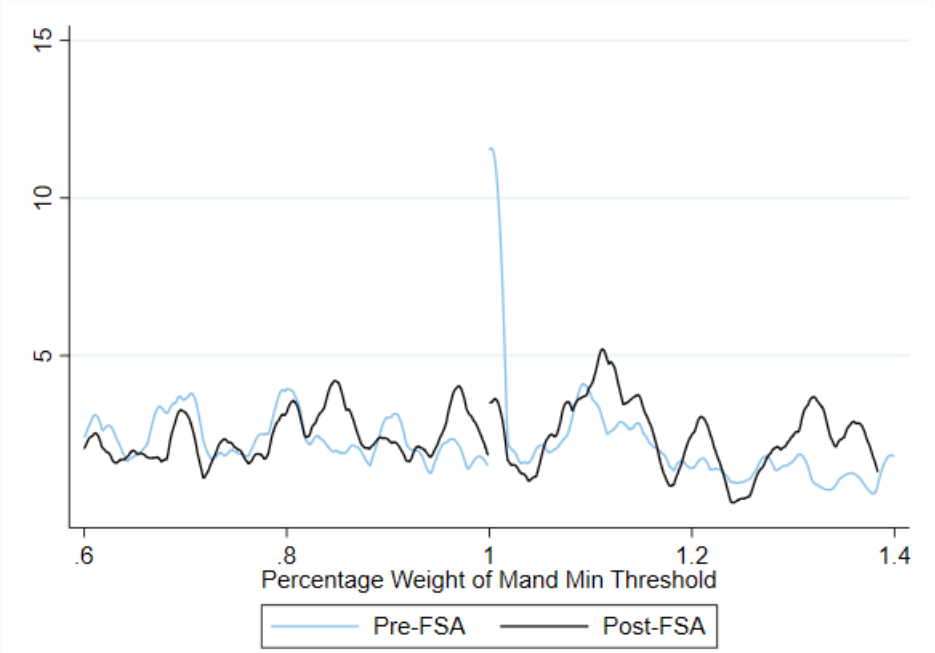


Figure A.7: Assessing Excess Mass for Groups with Decreased Bunching

(a) White Defendants, Conspiracy Circuits



(b) Minority Defendant Cases, Individual Circuits

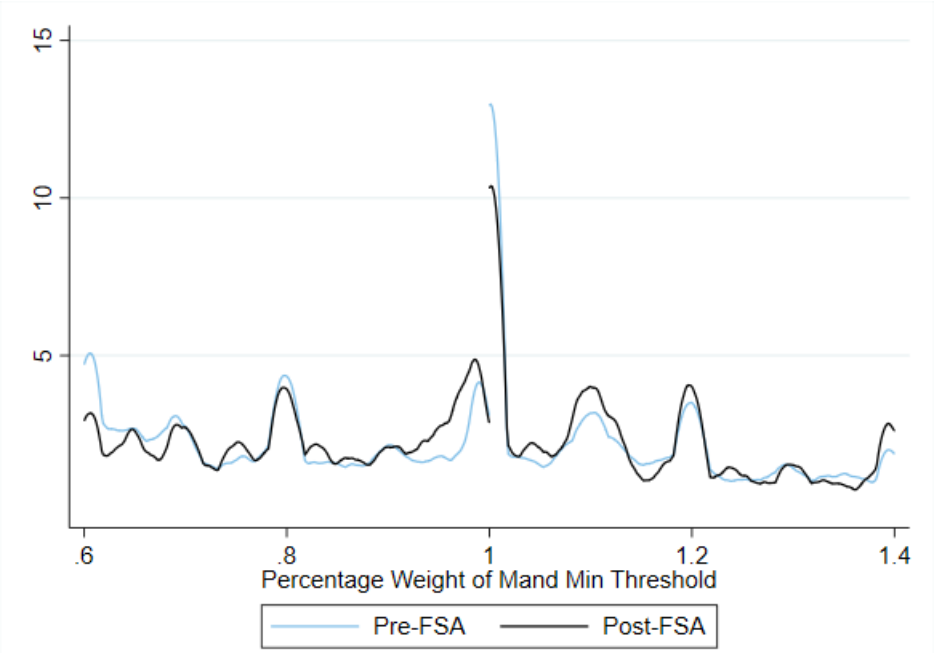


Figure A.8: Effects of FSA by Drug Type - Individual Circuits, Minority Defendants

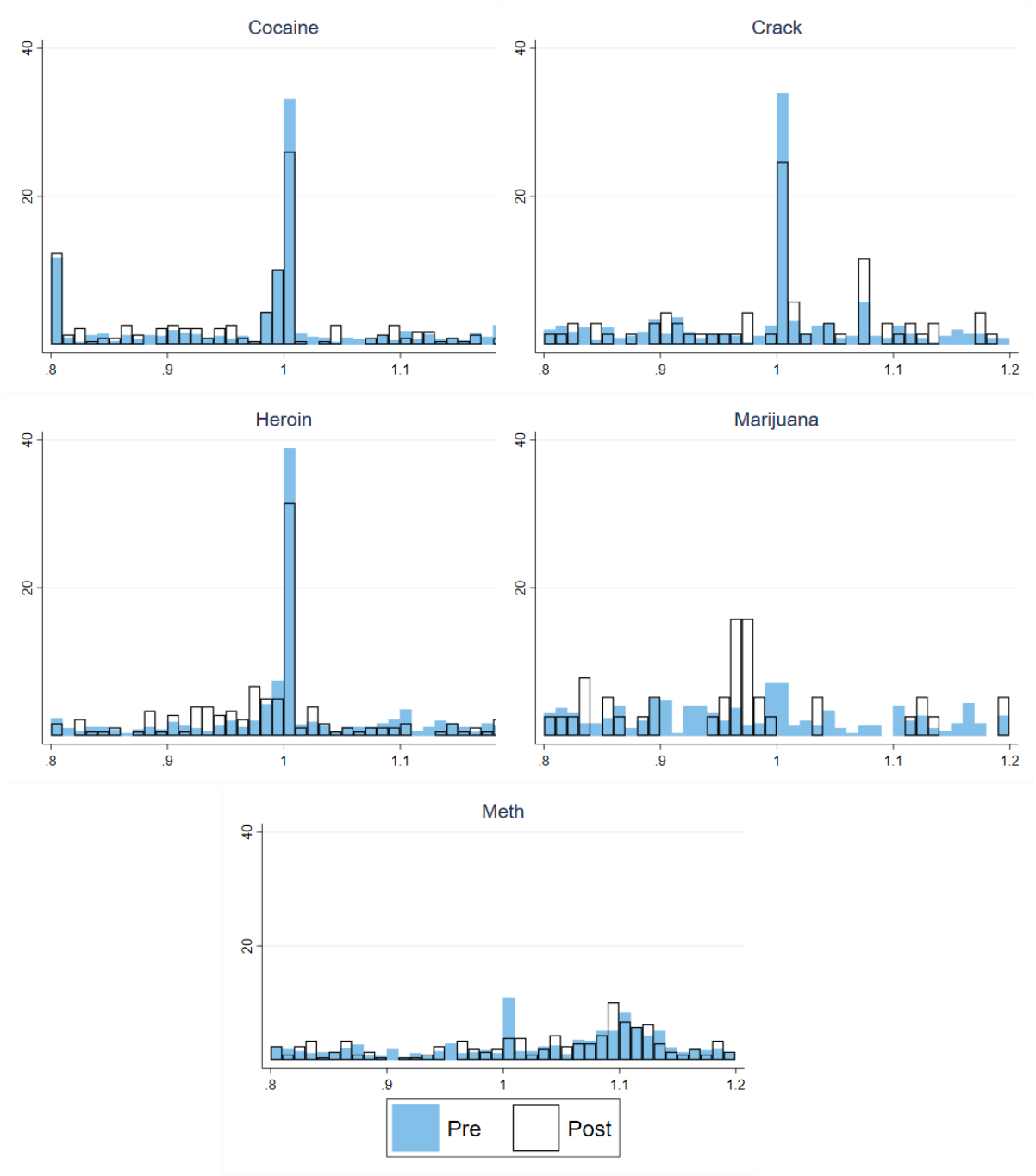


Figure A.9: Effects of FSA by Drug Type - Conspiracy Circuits, Minority Defendants

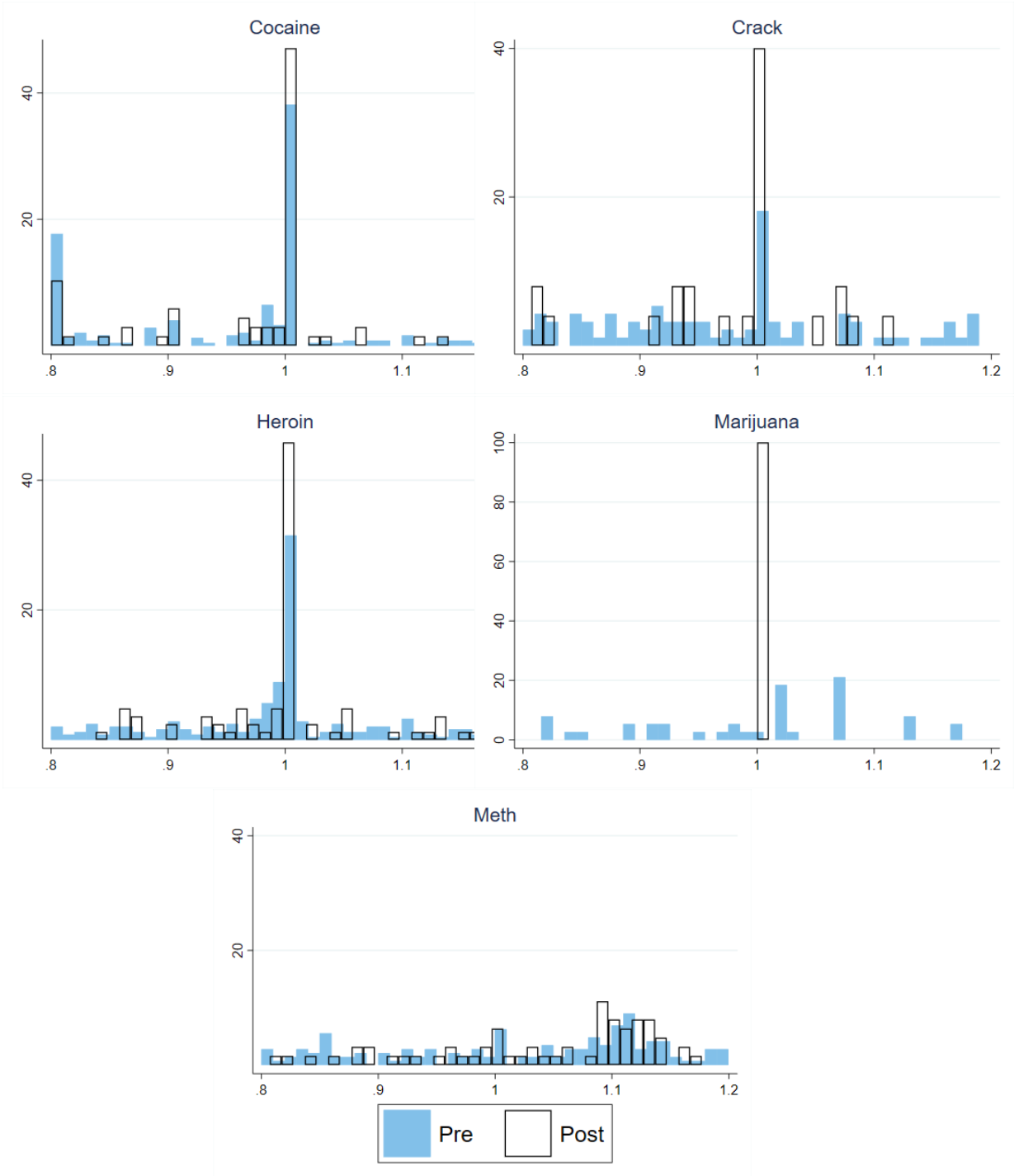
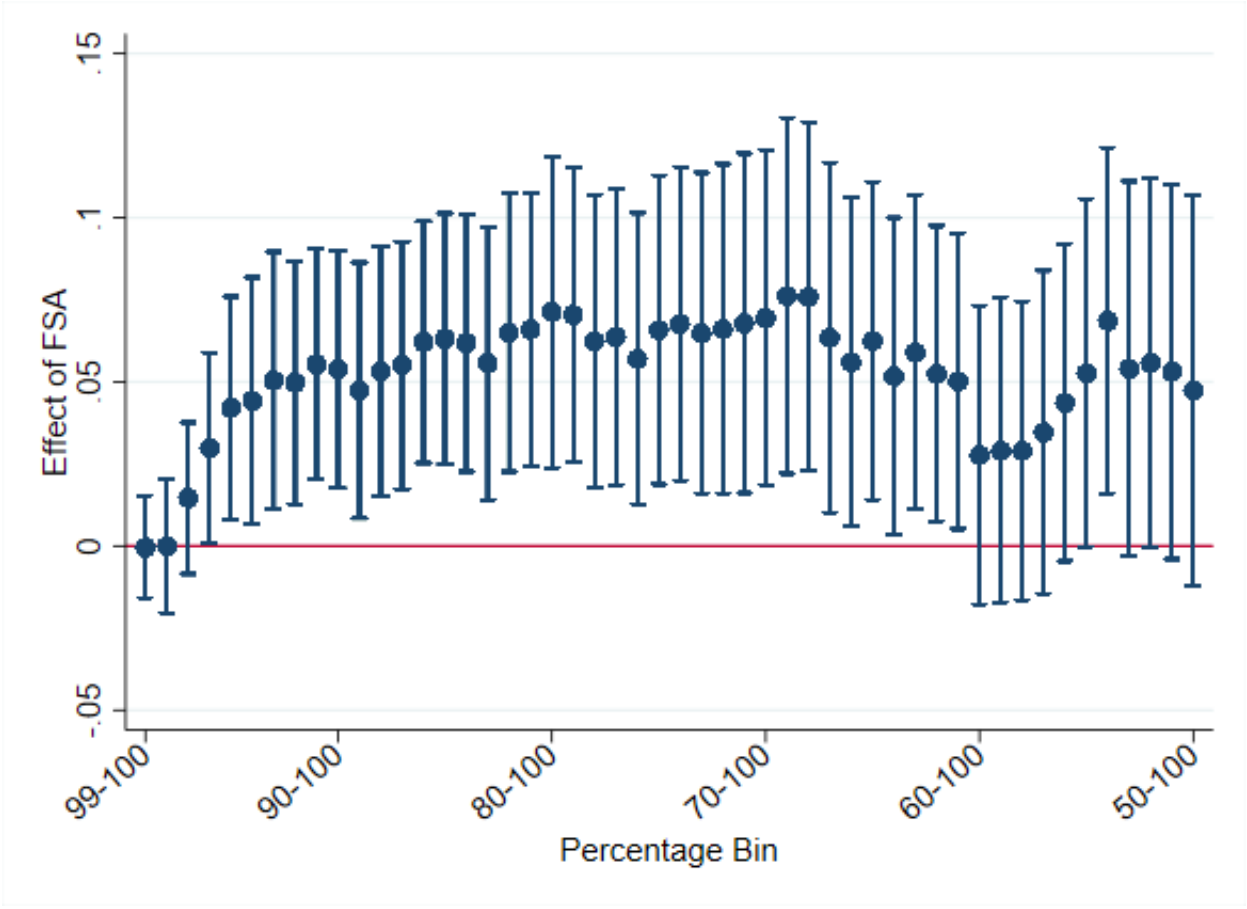


Figure A.10: Rounding Analysis Robustness Check - Individual Circuits, Minority Defendants



Graph gives the individual circuit-minority difference in difference coefficient for rounding direction analysis with 95 percent confidence bars with varying bin size of the dependent variable. The idea is to show the excess mass analysis is consistent regardless of how the dependent variable is defined. Each point adds an additional charging weight to the dependent variable. For example, the point at 90-100 is identical to the main results shown in Table A.9, column 1. The point after it changes the dependent variable to be a binary that includes all cases with weights 89 percent of the threshold weight up to, but not including, the bunching point.