Prosecutor Tradeoffs and Race: Evidence from a Circuit Split*

Spencer Cooper

June 26, 2024

Abstract

I consider whether changing the costs or benefits to prosecutor charging decisions impacts racial disparities in charges. Prosecutors face a tradeoff between expending resources to gather evidence and building a stronger case. I exploit a circuit split in charging rules to observe variation in the cost of evidence requirements. I compare how prosecutors differentially increase charges for racial minorities across each evidence cost regime by utilizing a bunching point in mandatory minimum charging of federal drug crimes. Using a difference-in-differences design, I find that when evidence costs are lowered, prosecutors disproportionately increase charges for Black defendants compared to White counterparts. I then consider how decreases to the expected sentence length for increased charges affects prosecutor choice and racial disparities using the 2018 First Step Act (FSA). I find that in circuits with high evidence costs, the FSA decreased the race gap in increased charges. But in circuits with low required evidence, the FSA increased this disparity. The results suggest that increasing evidence costs is an effective policy tool for reducing racial disparities in charges.

*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, Emily Leslie, Brigham Frandsen, Stephen Ross, and several anonymous US Attorneys and Assistant US Attorneys for helpful comments and insight.

It is well known that racial minorities are overrepresented in the US criminal justice system. According to a 2019 Bureau of Justice Statistics report, incarceration rates are around 10 times higher for Black individuals and 5 times higher for Hispanic individuals compared to Whites. Likewise, for federal crimes, sentences for Black defendants are about 10% higher than Whites for committing the same types of crimes (Rehavi & Starr 2014). An emerging literature on prosecutor discretion shows that one of the main channels contributing to racial disparities is prosecutor charging decisions (Rehavi & Starr 2014; Sloan 2022; Tuttle 2023).

In making charging decisions, prosecutors face a tradeoff between gathering evidence and building a case. Gathering evidence is costly and prolongs the litigation process, but may benefit prosecutors in two ways. First, additional evidence may increase the probability of conviction through plea deal or at trial. Second, additional evidence may allow prosecutors to increase the severity or number of charges above what they might otherwise charge a defendant with. If evidence costs are especially low or sentence length impacts to increasing charges are especially high, then prosecutors are far more likely to gather evidence and increase charges. Then if race disparities in charges are driven by prosecutor discretion, increasing evidence costs or decreasing punitive impacts of higher charges may reduce this race gap in punishment.

There are two main challenges in studying this evidence-charges tradeoff prosecutors face. First, researchers do not observe evidence gathering or charging decision discretion. Even if data were available on evidence submitted, it would be difficult to determine the costs to evidence gathering and impossible to know what charges the prosecutors were going to give before deciding for or against increase charges. I overcome this challenge by utilizing a bunching point in charging decisions created by mandatory minimum sentence laws for federal drug crimes as a rough measure of prosecutor's decision to increase charges. The second challenge is that it is difficult to observe variation in these tradeoffs. Evidence costs are typically unobserved by researchers and increased sentence lengths are only observed as outcomes endogenously affected by the prosecutor decision. I exploit a circuit split in evidentiary rules that create variation in evidence costs to increasing charges for prosecutors. The basic idea of my strategy is to compare prosecutor decisions to increase charges between

races under each evidence cost regime, using the bunching point to measure prosecutor choice and the circuit split to vary evidence costs.

This paper considers prosecutor decisions to increase charges for federal drug cases for the five most common drug types from 2000 to 2021 in each continental district. Specifically, I focus on cases local to the 10-year mandatory minimum eligibility threshold. Eligibility for mandatory minimum sentencing is determined by the charging drug weight. Each drug type has a different cutoff weight, so I normalize these to be in percentages of the threshold weight and limit my main study to cases within 50%-150% of the cutoff. Prosecutors have more bargaining power for cases eligible for higher mandatory minimums, meaning they have high incentive to get defendants into the eligible weight range. This discrete cutoff in eligibility creates a discontinuity in prosecutor incentives, leading to a bunching point in charging. Bunching arises because prosecutors are willing to engage in costly evidence gathering in order to find evidence of more drugs to add to a defendants weight to get them to the 10-year eligibility. Note that I am using mandatory minimum sentencing as a measure of prosecutor choice rather than providing a comprehensive analysis on the impacts of mandatory minimums.[1]

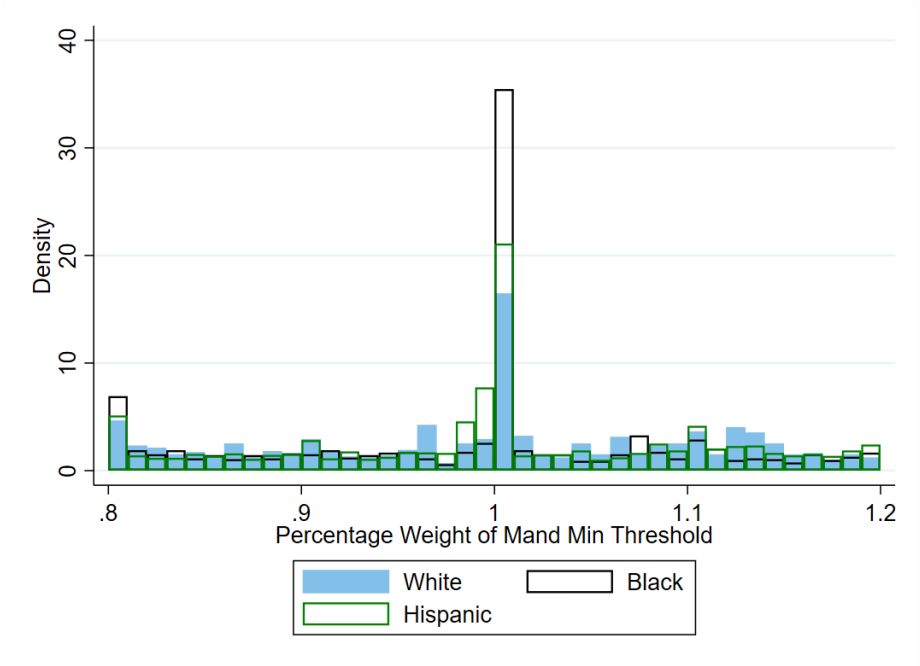
Figure 1 illustrates this bunching phenomenon for each race, aggregated across drug types. Note that Black and Hispanic cases have far more cases pushed to the bunching point compared to White cases.[2] This fact motivates the analysis for this paper; higher bunching for racial minorities indicates prosecutors choose to increase charges disproportionately for racial minorities. My analysis considers whether disparities can be reduced by increasing evidence costs or reducing sentence length impacts to increasing charges through mandatory minimums.

To estimate the effects of changing evidence costs on racial disparities, I employ a staggered difference-in-differences approach utilizing the timing of circuit court decisions to side with one cost regime or another. I find that when costs are increased or decreased, the total

[1]For analysis on the sentence length impacts of mandatory minimums sentencing, see the companion paper to this one Cooper (2023), as well as Bjerk (2017b).

[2]This racial disparity in bunching is driven most strongly cocaine and heroin cases but is still exists even when omitting them. Crack cases exhibit a similar disparity pattern, though not as large as one might expect given the focus it has received in past literature. The only drug type for which Whites receive higher bunching is marijuana, which makes up only 10% of all cases.

Figure 1: Bunching behavior by race



This graph contains three overlaying histograms, one for Black defendants (black and hollow), one for Hispanic defendants (green and hollow), and one for White defendants (blue and solid). These are not stacked vertically but are laying on top of one another, showing that racial minorities have stronger bunching compared to Whites.

number of cases bunched does not change. However, when evidence costs are lowered, prosecutors substitute away from bunching White defendants and increase bunching for Black ones. These results may be dampened by the fact that prior to a circuit decision, there is no status quo. This means some prosecutors may be completely unaffected by the decision if they were already operating similar to the regime implemented by the circuit decision. Regardless, the result indicates that when evidence costs are lowered, prosecutors increase charges for Black defendants more than Whites.

I then consider how decreasing the sentence length impact of mandatory minimums affects prosecutors' decisions to increase charges. I utilize the 2018 First Step Act (FSA) as a plausibly exogenous decrease in the expected sentence length of a mandatory minimum charge 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. Furthermore, in circuits with low costs to manipulation, bunching is significantly reduced for

White defendants but not for minority ones. Taken together, the results imply that evidence costs for Black and Hispanic defendants are lower than for White counterparts. This means that increasing evidence costs or decreasing sentence impacts are effective tools in decreasing racial disparities, but only if cost increases and sentence decreases are sufficiently large.

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 2023; 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 2023).

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 (2023).

The remainder of the paper is organized as follows. Section II provides more details about

how prosecutors increase charges and the circuit split. Section III presents a conceptual model of prosecutor manipulation choice. Section IV gives details on data and the empirical strategies employed in my main analyses. I then present the results and robustness checks in Section V. Section VI gives a discussion on the interpretation of the results and the policy implications. Section VII concludes the paper.

II Background

II. A. How Does Bunching Occur?

Bunching could be the result of prosecutors increasing charging weights or decreasing them through some form of leniency or plea bargaining mechanisms. Bjerk (2017b) suggests that bunching likely occurs from downward manipulation while Tuttle (2023) provides evidence that cases have weights increased. In this paper, I assume (and show evidence) that manipulation is a result of prosecutors increasing the charging weight rather than decreasing it.

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.

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."[3] For drug trafficking charges, conspiracy implies a defendant was planning or attempting to transport or distribute drugs, often working with others to do so.

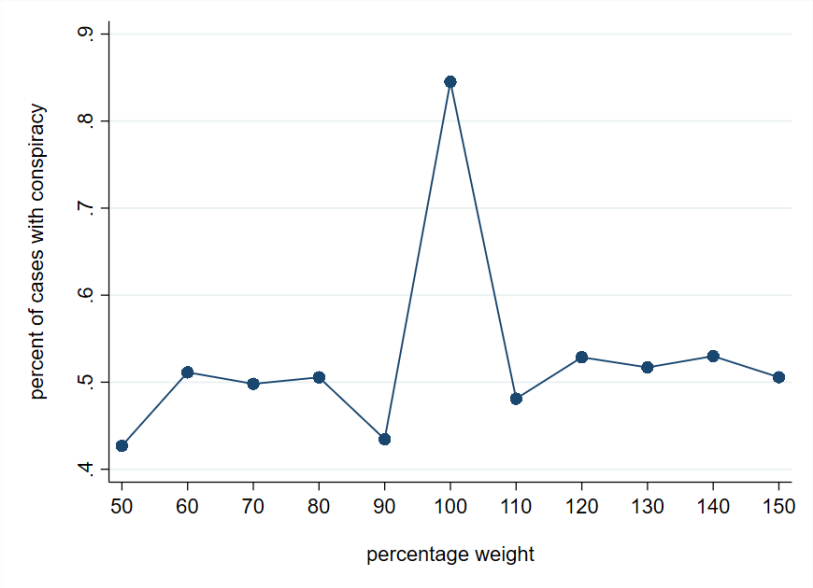
Conspiracy charges are a key method whereby charging weight can be manipulated upwards. For all drug trafficking cases, drug quantity may be determined either from seized quantity or by approximation. Approximation may be based on several factors including cash seized, testimonials, inputs and equipment, 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 (2023) 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 2 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.[4] 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.

[3] 21 U.S.C. § 846.

[4] 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 2: 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.

II. B. 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 within each circuit:

1. **The conspiracy approach** - 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).

2. **The individual approach** - this 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.[5] 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.[6]

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 3 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,[7] and the latest being in June of 2018.[8] 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.

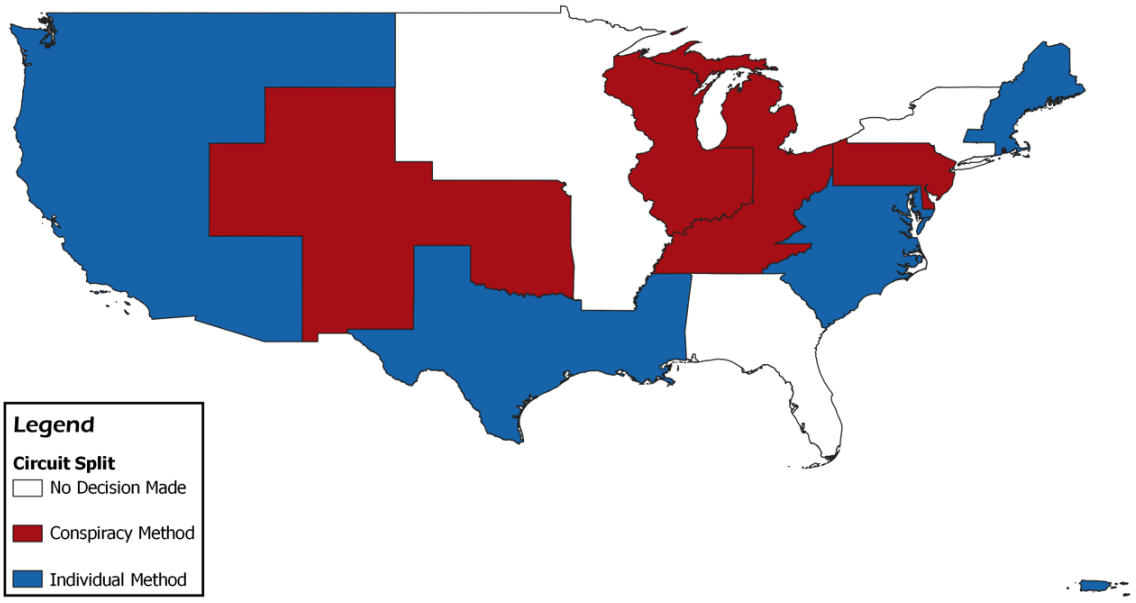
[5]Conspiracy clearly still has an effect on manipulation; when comparing conspiracy charges by weight, the strong bunching seen in Figure 2 holds when only including non-conspiracy approach circuits.

[6]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).

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

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

Figure 3: Conspiracy Drug Weight Approach by Circuit



II. C. FSA Details

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.[9]

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

[9]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.

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.

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 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 and Empirical Strategy

IV. A. Data

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.[10] Data

[10] 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."

is restricted to the five most prevalent substances subject to mandatory minimum sentencing: powder cocaine, crack, heroin, methamphetamine, and marijuana.[11] 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 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.2 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

Effectively, it raised the evidentiary burden of proof for drug weight charging, thereby affecting prosecutor incentives to manipulate drug weights.

[11]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

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 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.[12] 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.

IV. B. Empirical Strategy

Changing Evidence Costs - the Circuit Split Decision

I use the timing of circuit court decisions as plausibly exogenous changes to evidence costs for increasing charges against a defendant. The timing of these decisions happened over a wide period of time, with the earliest decision occurring in March of 2003 and the most recent decision being made in June of 2018. To compare decisions made across this wide time span I adopt a staggered difference-in-differences model. I consider three years before and after each decision centered on treatment timing and consider effects separately for circuits that adopt the conspiracy method versus those that adopt the individual method.

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

Currently, this strategy only considers differences in racial composition of bunching within circuit type and does not use case variation in circuits without a decision made.

The model is as follows:

$$\begin{aligned} Bunched_{idtm} = \alpha + \beta_1 Black_i + \beta_2 Hispanic_i + \beta_3 post\_decision_{idtm} \\ + \beta_4 Black_i \times post\_decision_{idtm} + \beta_5 Hispanic_i \times post\_decision_{idtm} \\ + X_i\gamma + \eta_d + \epsilon_{idtm} \quad (2) \end{aligned}$$

$Bunched_{idtm}$ is a binary variable for case i being charged at the threshold weight in district d during year t and month m. It serves as a rough measure of the prosecutor's decision to increase charges. $post\_decision_{idtm}$ is a binary measure that equals one if case i is sentenced after the timing of the circuit court decision for the circuit it operates in. White defendants are the omitted group in this model, meaning β_3 captures changes in the percent of White cases at the bunching point, while $\beta_3 + \beta_4$ and $\beta_3 + \beta_5$ give the effects of changing evidence costs on Black and Hispanic cases, respectively.

I consider this model under a series of different control schemes. In specifications controlling for covariates, I include vector X_i which contains drug type fixed effects, criminal history points, age and age squared, sex, a binary for college attainment, and a binary for being an illegal alien. I also include district fixed effects to control for geographic shocks. I also include a specification in which I saturate all controls by circuit. This is equivalent to performing the analysis for each circuit separately and aggregating the difference-in-difference effects.

Decreasing Sentence Impacts - the FSA

To compare the effects of decreasing expected sentence lengths on prosecutor decisions, I employ a difference-in-differences approach to the bunching analysis utilizing the timing of the FSA. The following model captures the differential impact of the FSA on racial minorities

compared to White counterparts:

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

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 circuit split model. The main differences are that I include US Attorney and month-of-year fixed effects, I include annual racial population measures, represented as P_{dt}, 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.[13] 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.[14] 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.

[13]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.

[14]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.

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.

Threats to Identification

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 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. This does not appear to be the case as enhancement rates are fairly similar across race for cases within the 50%-150% weight window. 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. 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.

V Results

V. A. Changing Evidence Costs Effects

I first run the model in Equation 2 without the difference-in-difference estimators to naively capture the difference in overall percent of cases being charged at the threshold weight after the circuit court decision. I find that in both individual and conspiracy circuits, decisions did not have a distinguishable effect on the proportion of cases being bunched. However, it is possible that while the same proportion of cases is bunched, the racial composition of these cases changes after evidence costs change.

Table 1 displays the results for circuits that decided towards the conspiracy approach.

Each specification gives a different control scheme, with the last column providing the fully saturated by circuit model. The results tell a consistent story across each specification; bunching for Black defendants increases while for White cases it decreases. Bunching for Hispanic defendants remains unchanged. This suggests that as evidence costs decrease and prosecutors have more ability to increase charges, they substitute away from White cases and towards Black ones. The fact the White case bunching decreases is surprising, but may point to capacity constraints if the composition in cases itself changes after the decision.

Table 1: Circuit Decision DID for Conspiracy Method Circuits

	(1)	(2)	(3)	(4)
Black	-0.0447*	-0.00906	-0.00657	-0.0157
	(0.0248)	(0.0239)	(0.0214)	(0.0203)
Hispanic	-0.0293	-0.0476	-0.0257	-0.0297
	(0.0272)	(0.0287)	(0.0259)	(0.0236)
split	-0.0393	-0.0409*	-0.0449**	-0.0403*
	(0.0242)	(0.0227)	(0.0217)	(0.0208)
Black×split	0.0566**	0.0534**	0.0603**	0.0567**
	(0.0266)	(0.0248)	(0.0241)	(0.0241)
Hispanic×split	0.0318	0.0368	0.0417	0.0369
	(0.0307)	(0.0287)	(0.0268)	(0.0255)
Covariates		X	X	X
District FE			X	X
Saturated Model				X
Observations	3947	3947	3947	3947

Standard errors clustered at the district level in parentheses

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

The results for the individual circuit decisions are displayed in Table A.3. Here, results are consistently null. This finding is somewhat expected - three of the five circuits that switched to the individual approach did so after the ruling of *Alleyne v. United States*. While *Alleyne* did not explicitly create precedent on evidentiary rules for drug trafficking conspiracy charges, it broadly suggests that the individual approach should be followed.[15] Thus, decisions to switch may have had little impact because these circuit were already

[15]This may seem like there should be no difference between the circuit approaches if *Alleyne* suggests the individual approach should be followed. However, conspiracy circuits have not changed their precedent and in some cases, have explicitly confirmed that the conspiracy method would still be precedent within their circuit despite the *Alleyne* ruling.

largely following the individual approach.

While the conspiracy method switch appears to have some impact on bunching behavior, these results may undersell the effect evidentiary costs can have on prosecutor discretion. Before a circuit decision is made, it is unclear what the charging practice status quo is within each circuit. Thus, some circuits may formally change to the conspiracy or individual approach while many cases were already carried out as if that approach were the standard. Thus, if anything the effect of evidence costs on prosecutors' decisions to increase charges is likely higher than what is reported above.

V. B. Decreasing Sentence Lengths Effects

The FSA caused a large, immediate drop in bunching at the high threshold. Figure A.3 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.[16] 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 4 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

[16]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.4 for Black and Hispanic effects separately.

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 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 4. 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.

Figure 4: Effects of FSA by Circuit Group and Race

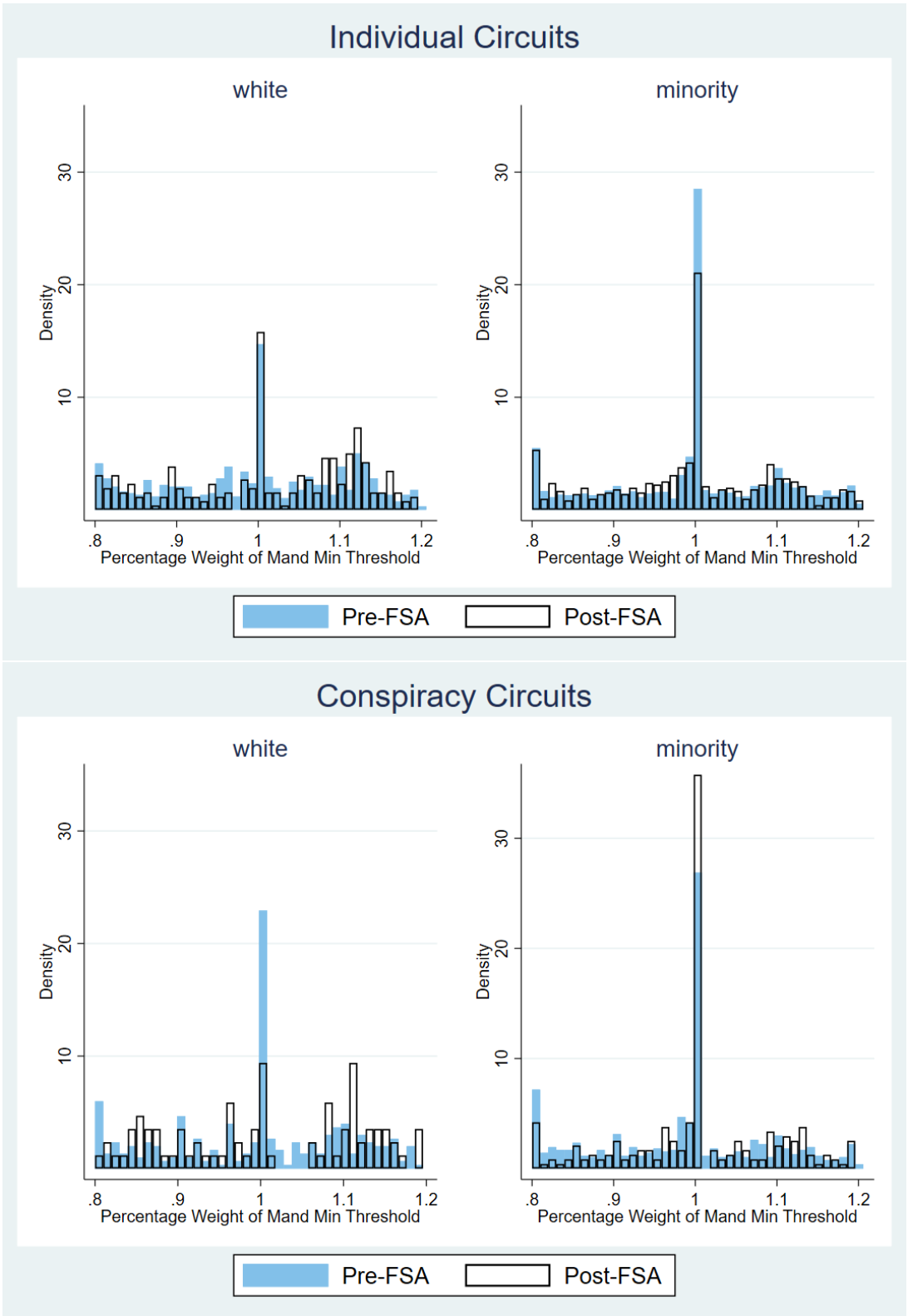


Table 2: Changes in Bunching by Race and Circuit Type

	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)
mean white pre-FSA bunching	0.0542	0.0542	0.0997	0.0997
mean minority pre-FSA bunching	0.1127	0.1127	0.1123	0.1123
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.5 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.5a 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.5b 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 II, I provide evidence that bunching is a result of upward manipulation

indicated by the strong correlation with conspiracy charges. 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.

Figure A.6 presents kernel density graphs showing the pre and post distributions for the groups with bunching decreases. While Figure A.6a is noisy and uninformative, A.6b 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. Table A.5 presents the difference-in-difference results for the model using a binary measure of a case being just below the threshold for individual circuits. 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.

VI Discussion

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.2 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^*.

Taken together, the FSA results across each of the four groups along with the effects of circuit court decisions describe a pattern of prosecutor behavior. As evidence costs increase (or sentence lengths decrease), prosecutors choose to bunch less cases. They first stop pursuing White cases, and continue bunching minority cases. Then at some point, raising evidence

costs begins to decrease bunching for minorities. This means that on average, minority defendants have a lower evidence cost than White defendants. This is why high bunching remains for minorities in conspiracy circuits - the costs to bunching are so low prosecutors still prefer to increase charges even after sentence lengths decrease. This result is consistent with the interpretation from the conceptual model.

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.

I first vary the definition 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. I then consider the FSA results for a larger samples of cases. These results are presented in Table A.6. 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 FSA histograms for minority defendants in individual circuits by each of the five drug types in Figure A.7. 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 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.[17] 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

Prosecutors face a tradeoff between expending resources to gather evidence and building a case. Prosecutors may wish to increase charges at the cost of more evidence. Little is known about this tradeoff decision because researchers do not observe evidence gathering or charging discretion. I utilize a bunching point in charges created by mandatory minimum eligibility as a measure of prosecutors choice to increase charges. I note that large racial disparities exist in bunching across the five drug types studied. Exploiting a circuit split in evidence costs, I consider how prosecutors change their decisions to increase charges when costs to evidence are increased or expected sentence length benefits decrease. I specifically consider how these policy interventions impact race gaps in charging decisions.

I first consider circuit court decisions to increase or decrease the evidence costs needed to increase charges. I find that when evidence costs are decreased, prosecutors substitute away

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

from bunching White defendants and increase bunching for Black defendants. Decreasing evidence costs appears to have no effect on the number or composition of cases bunched, though this may be due to those circuits already having high evidence costs to begin with. I then consider the effects of decreasing sentence length impacts. This is done by assessing the bunching responses to the First Step Act across each circuit type. I find this policy effectively reduced the race gap in circuits with high costs to bunching. Specifically, in these high cost circuits, bunching decreased dramatically for minority defendants White bunching exhibited no changes. But the race gap actually increased following the FSA in circuits with low costs to bunching. Taken together, these results imply that evidence costs are lower for Black and Hispanic defendants compared to White counterparts.

On a broader scale this paper suggests that prosecutors are sensitive to changes regarding this evidence-sentence tradeoff. If reducing racial disparities is a goal for policymakers, the policy recommendation from this paper is to increase evidence costs or reduce sentence length impacts. In the specific context of mandatory minimums, prosecutors have discretion as to which types of defendants to pursue for higher charges. This paper gives evidence that they disproportionately pursue higher charges for racial minorities for committing the same types of crimes as White counterparts. Utilizing evidence costs or sentence length benefits through policy appears to be an effective way to reduce this race gap.

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

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.

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.

Cooper, S. (2023). *How much do mandatory minimums matter?* (Working Paper). University of Arizona.

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*.

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.

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.

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.

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.

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.

Sloan, C. (2022). *Do prosecutor and defendant race pairings matter? evidence from random assignment* (Working paper).

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. (2023). *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.[18] 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.[19] 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

[18] 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.

[19] 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.[20] 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.

[20]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: 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.3: Circuit Decision DID for Individual Method Circuits

	(1)	(2)	(3)	(4)
Black	0.0312*	0.0452**	0.0234	0.0278
	(0.0179)	(0.0199)	(0.0167)	(0.0169)
Hispanic	0.00758	-0.0187	0.00105	0.000399
	(0.0136)	(0.0119)	(0.0129)	(0.0128)
split	0.0214	0.0220	0.0230	0.0225
	(0.0209)	(0.0185)	(0.0193)	(0.0176)
Black×split	-0.0111	-0.0110	-0.0120	-0.0142
	(0.0224)	(0.0203)	(0.0231)	(0.0213)
Hispanic×split	-0.0174	-0.0209	-0.0211	-0.0211
	(0.0171)	(0.0165)	(0.0179)	(0.0161)
Covariates		X	X	X
District FE			X	X
Saturated Model				X
Observations	6308	6308	6308	6276

Standard errors clustered at the district level in parentheses

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

Table A.4: 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.5: 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.6: 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

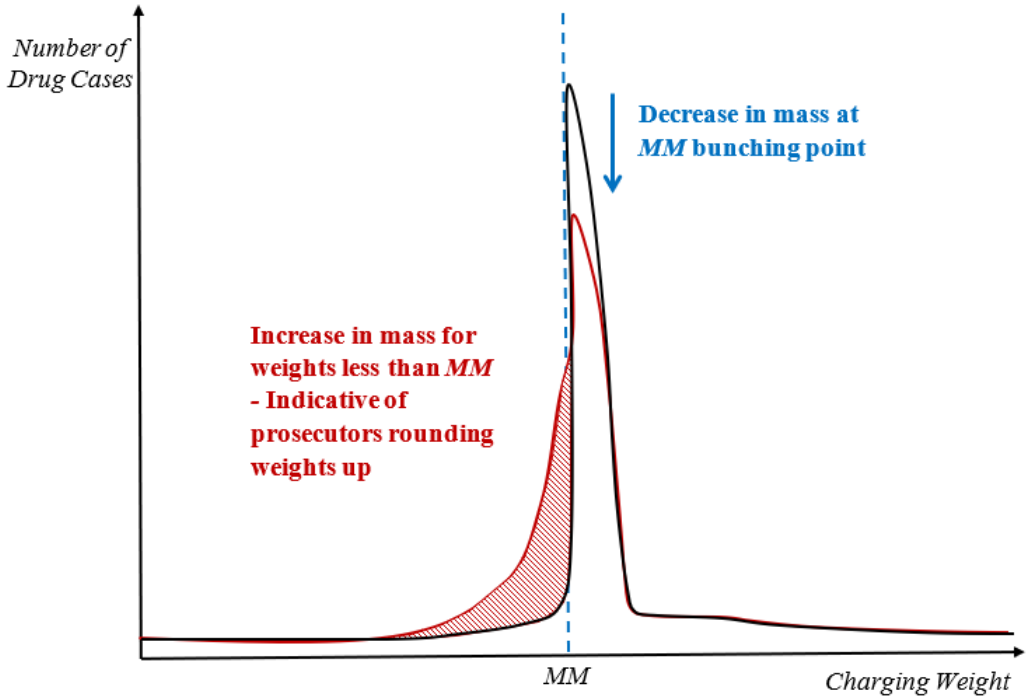
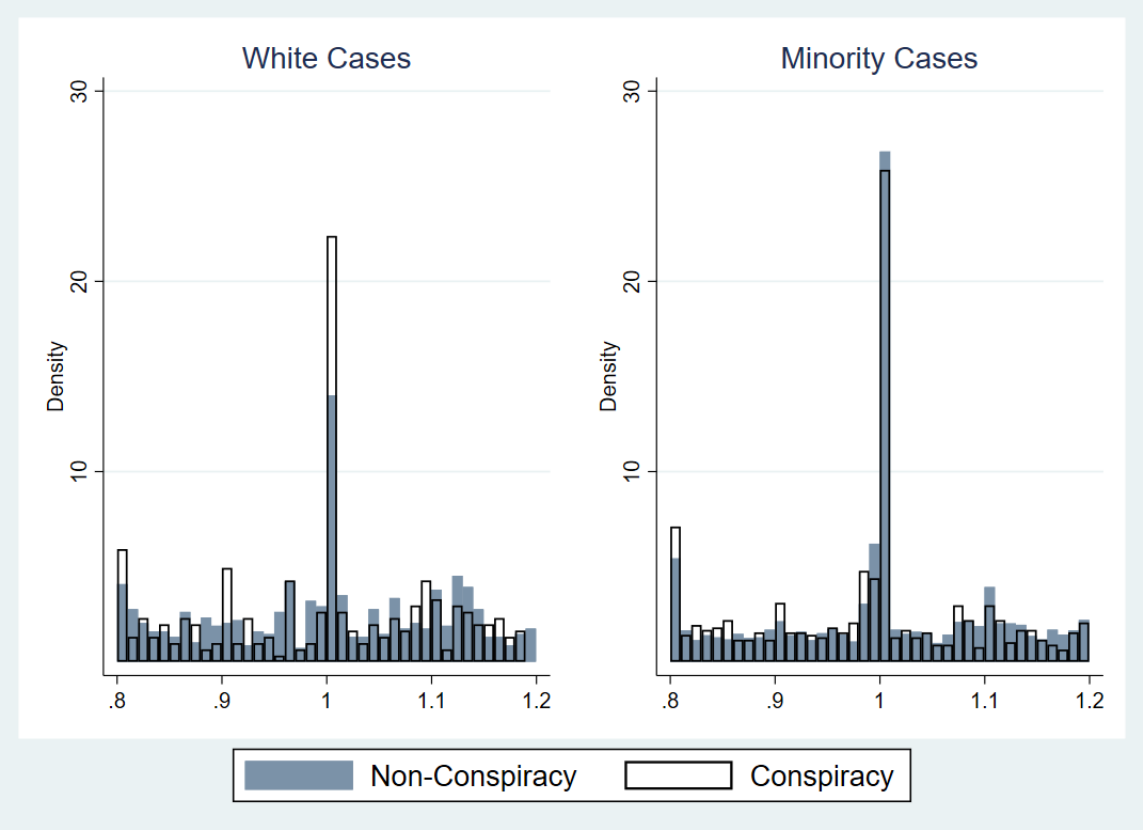


Figure A.2: Pre-FSA Bunching Patterns



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.3: Charged Drug Weight Pre vs Post FSA

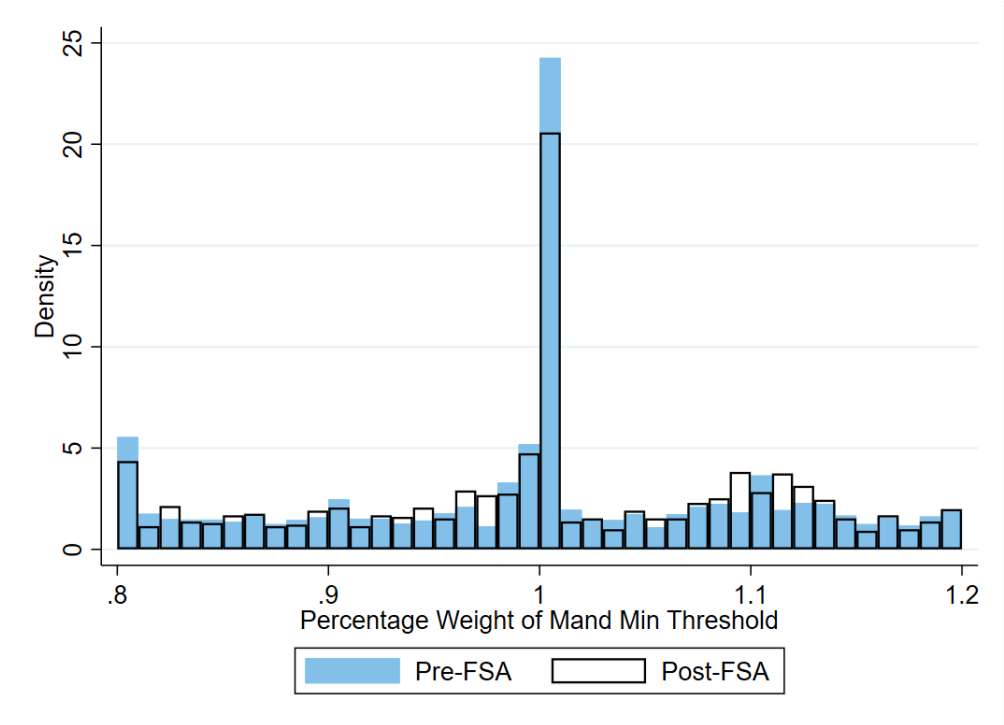


Figure presents a histogram of charge weights for all cases in the sample before the FSA was enacted versus all cases that occurred after.

Figure A.4: Circuit Split DID Trends for Black Cases in Conspiracy Circuits

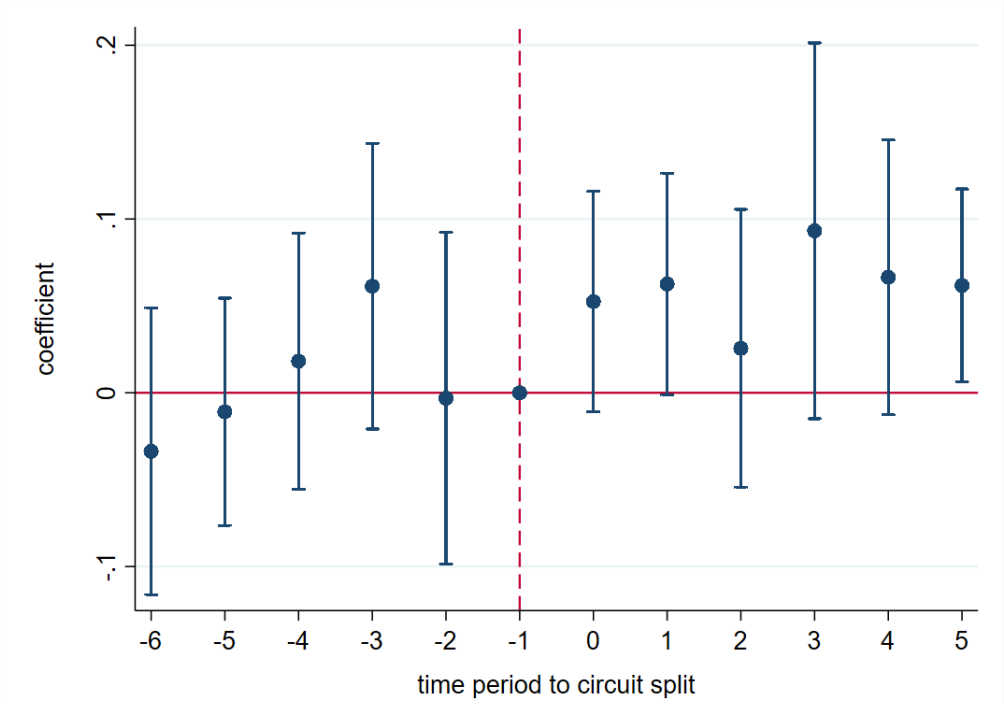
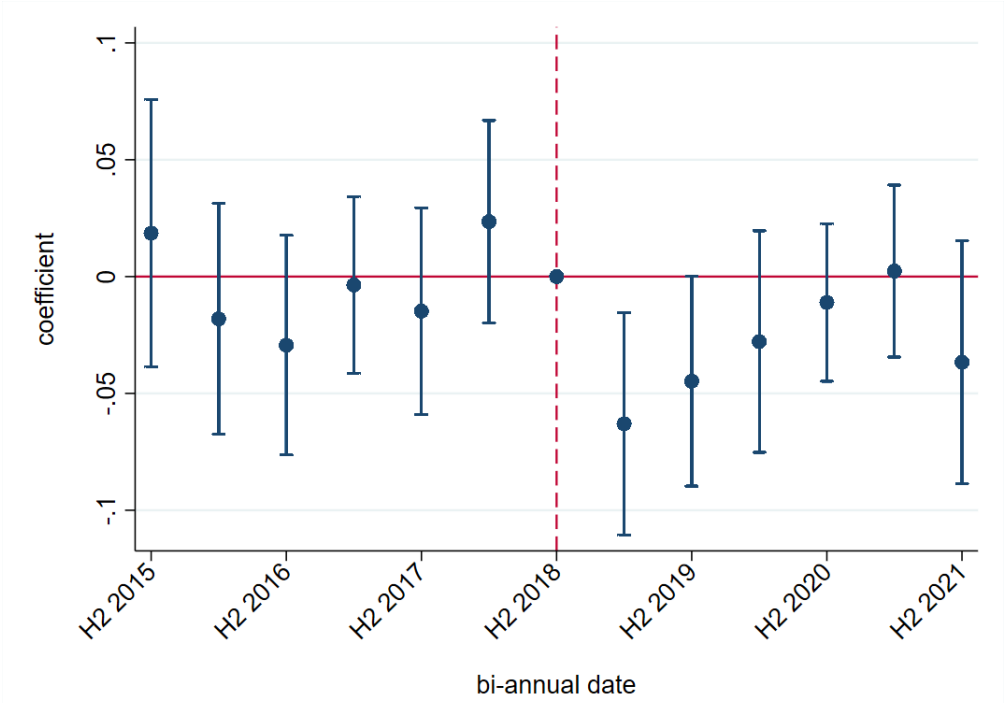


Figure presents the DID coefficient for Black cases compared to Whites for each time period relative to the circuit decision

Figure A.5: FSA DID Trends

(a) Individual Circuits



(b) Conspiracy Circuits

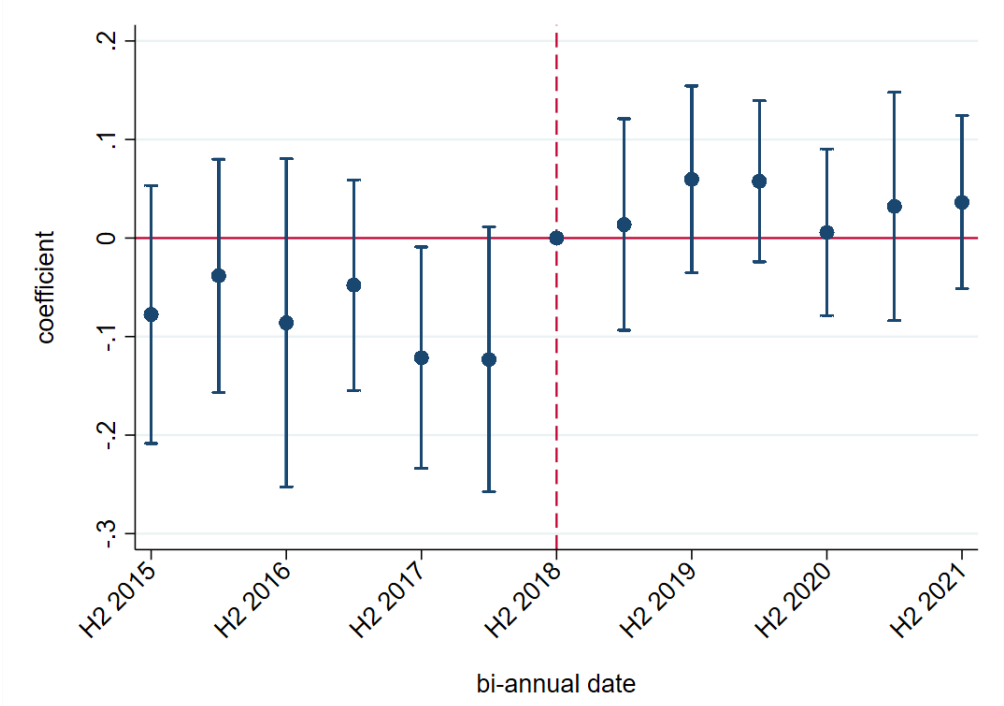
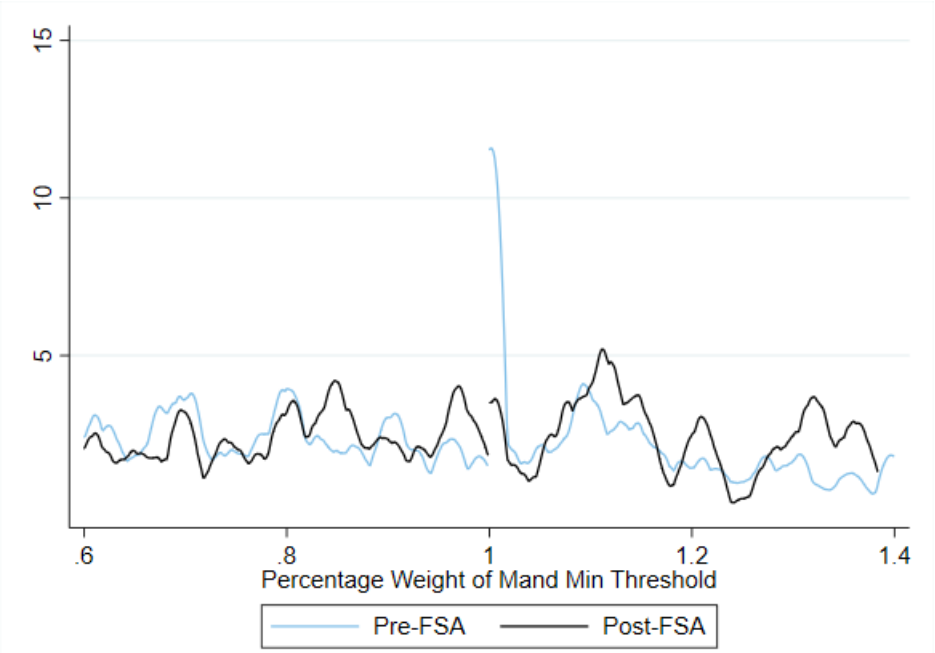


Figure A.6: Assessing Excess Mass for Groups with Decreased Bunching

(a) White Defendants, Conspiracy Circuits



(b) Minority Defendant Cases, Individual Circuits

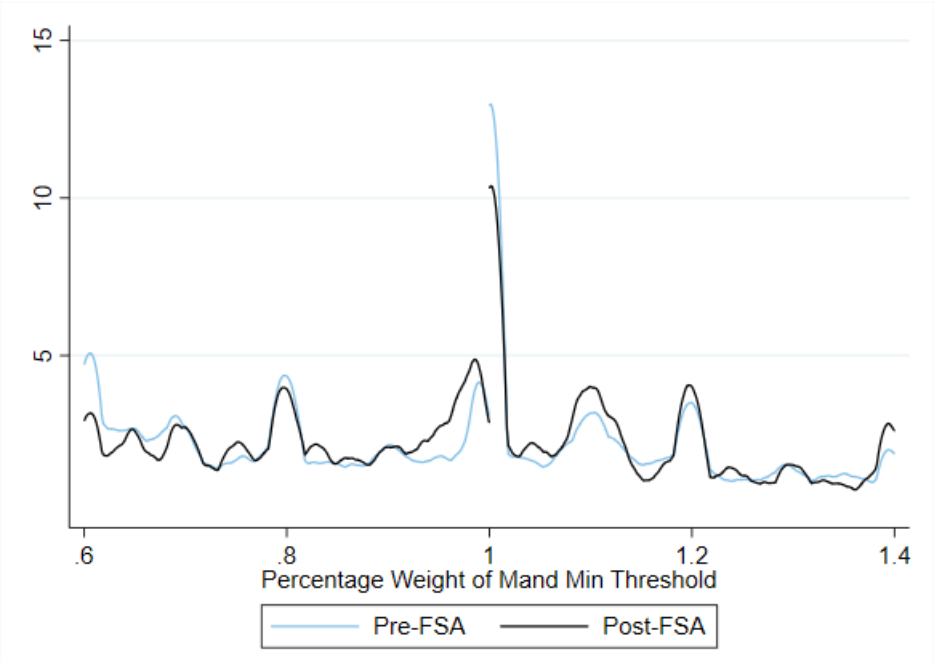


Figure A.7: Effects of FSA by Drug Type - Individual Circuits, Minority Defendants