Prosecutor Tradeoffs and Race: Evidence from a Circuit Split*

Spencer Cooper

July 14, 2025

Abstract

Prosecutors face a tradeoff between gathering evidence and pursuing charges. I examine how differences in evidentiary costs affect racial disparities in charging decisions by exploiting a circuit split in federal drug sentencing rules. Comparing high- and low-cost regimes, I find that mandatory minimum policies reduce preexisting race gaps by lowering penalties for minorities. However, when no gap exists initially, benefits accrue mainly to White defendants, creating disparities. High-cost regimes also show higher pre-policy punishment levels. The findings suggest raising evidence costs can reduce existing disparities but may introduce new ones when race gaps are absent.

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 about 10 times higher for Black individuals and 5 times higher for Hispanic individuals compared to Whites. Likewise, sentence lengths for federal convictions are about 10% higher for Black defendants than for Whites counterparts (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 2025).

When making charging decisions, prosecutors face a tradeoff between collecting additional evidence and advancing the case. Collecting evidence is costly and can lengthen the litigation process, but it can benefit prosecutors in two key ways. First, it can increase the likelihood of securing a conviction, either through a plea agreement or at trial. Second, it may allow prosecutors to bring more serious or numerous charges than they otherwise could. When evidence costs are low or when higher charges substantially lengthen potential sentences, prosecutors are more inclined to invest in gathering additional evidence and increasing charges.

If racial disparities in charging stem from prosecutorial discretion, then raising the cost of evidence collection or reducing the punitive consequences of higher charges may help narrow these disparities. This paper examines how differences in evidentiary costs contribute to racial disparities in prosecutors' charging decisions.

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 increased charges. To overcome this challenge, I consider the specific context of mandatory minimum (MM) sentencing for drug crimes. MM sentencing provides a setting in which prosecutors exercise wide discretion over charging decisions and where decisions to increase charges can be proxied for by exploiting features of the MM eligibility rules.

The second challenge is that it is difficult to observe variation in evidentiary costs. Ideally, the researcher would observe whether prosecutors increase charges under a schedule of

1

different evidentiary costs. This is difficult not only because evidence costs are unobserved, but also because liability rules are often homogeneous within a given court system. To overcome this, I exploit a circuit split which creates variation in the rules for liability burden for drug trafficking cases. Essentially, some circuit courts explicitly ruled to have more strict or less strict rules regarding evidence needed for MM eligibility. While each case will carry different costs to increasing charges depending on the crime and defendant, the circuit split creates a a direct change in the baseline cost for any given case. Taken together, I address these challenges by using MM charging measures to proxy for prosecutor choice and the circuit split to vary evidence costs.
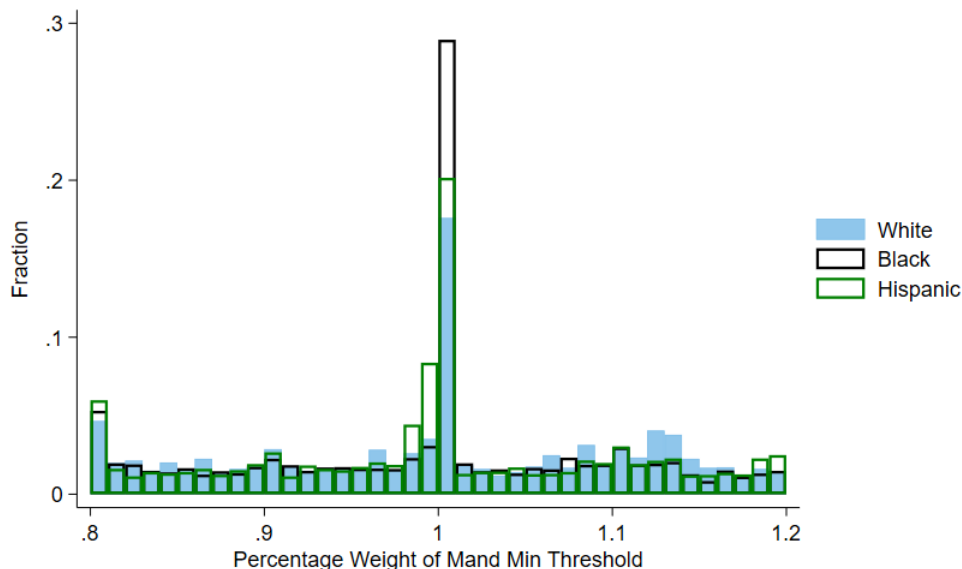
This paper considers prosecutor decisions to increase charges for federal drug cases for the five most common drug types from 2008 to 2021 in each continental district. Specifically, I focus on cases local to the 10-year MM eligibility threshold where prosecutors are most likely to have discretion over whether to seek MM charges or not. Eligibility for MM sentencing is primarily 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. However, prosecutors also choose whether to pursue the application of MM charges. This is another dimension by which prosecutors can choose to expend resources to gather evidence in order to secure a higher sentence length. These two measures serve as the key dependent variables in my analysis. Note that I am using MM sentencing as a measure of prosecutor choice rather than providing a comprehensive analysis on the impacts of mandatory minimums.[1]

MM sentencing also serves as an ideal context to study race disparities. MM sentencing

---

[1]For analysis on the sentence length impacts of MM sentencing, see the companion paper to this one Cooper (2025), as well as Bjerk (2017b).

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. Histograms represent all cases with weights 80% - 120% of the threshold weight from 2008 to 2020.

has been historically linked with racial injustice due to discrepancies in eligibility by drug type, and empirical research has provided evidence that racial minorities are disproportionately punished by MM sentencing (Cooper 2025; Rehavi & Starr 2014; Tuttle 2025). 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.

To estimate these effects, I consider the impact of three major MM policies on race disparities and compare responses across these two evidence-cost regimes. These three policies are the Fair Sentencing Act of 2010 (Fair 2010), the Drugs Minus 2 Amendment of 2014 (Minus-2), and the First Step Act of 2018 (FSA). Each of these provide a plausibly exoge-

---

[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.

nous shock to prosecutor incentives - Fair-2010 increases costs to attaining a MM charge for crack cases, Minus-2 decreases baseline sentence lengths for most trafficking cases, and FSA decreases sentence length increases for a MM charge.

I compare bunching levels and MM charge applications before and after each policy across race and circuit type, using difference-in-difference and triple difference designs. I find that under each regime, policies are effective in reducing pre-policy race gaps in charging behavior. However, if no race gaps exist before the policy, race gaps are induced as benefits of the policies are localized to White defendants. Thus, policies appear to provide benefits to one racial group, with White defendants moving first. Furthermore, circuits operating under the low-cost regime lag behind high-cost ones. These findings suggest that evidentiary costs are lower for White defendants compared to minority counterparts, and adopting a unified high-cost evidentiary practice for all circuits would likely decrease race gaps in high-cost regime circuits.

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 2025; 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. (2023) 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 2025).

The remainder of the paper is organized as follows. Section II provides more details about how prosecutors increase charges, the circuit split, and each policy. Section III gives details on data and the empirical strategies employed in my main analyses. I then present

the results and robustness checks in Section IV. Section V concludes the paper.

II Background

II. A.   MM Eligibility and Application

Convictions carrying a mandatory minimum require sentence lengths to be greater than or equal to some specified amount. In federal drug cases, eligibility for a MM sentence is mainly determined by the quantity of drugs a defendant is charged with.[3] There is a lower MM charge, which carries a 5-year minimum, and a higher one, which carries a 10-year minimum, and each drug type has different weight eligibility thresholds. In this paper, I focus on 10-year MM charges as this higher threshold has distinct bunching, higher sentence length punishment, and still applies to many cases (over 60%).

Not every case that is eligible for MM sentencing has a charge that carries the MM application. Prosecutors must decide whether to apply MM charges to cases that are eligible. This decision is impacted by several factors including the strength and amount of evidence prosecutors have, and negotiations in plea bargaining. Cases that are eligible for MM sentencing but do not have the MM application still receive higher sentence lengths than non-eligible cases. In the companion paper to this one, Cooper (2025) estimates 10-year MM eligibility increases sentence length by about 10 months. But applied MM charges also greatly increase sentence lengths - in my data, I find that average sentence lengths for cases without MM application have an average sentence of around 70 months, while cases with application have an average sentence of around 105 months. Thus, if prosecutors wish to achieve a higher sentence length, they may expend resources across one or both of these margins: securing 10-MM eligibility and choosing to apply MM charges. I measure MM application directly, but I cannot directly observe whether prosecutors engage in behavior that increases drug weight. However, I can proxy for this manipulation behavior by exploiting a bunching point in the distribution of drug weights.

---

[3]Other factors such as having a high criminal history level or using a firearm during the crime can increase eligibility from one level to another (for example, a 5-year to a 10-year MM charge), but baseline eligibility is determined by drug weight.

How Does Bunching Occur?

As shown in Figure 1, the strict eligibility cutoff for MM sentencing creates a bunching point in the drug weight distribution right at the threshold weight. Bunching may be the result of prosecutors increasing charging weights from lower parts of the distribution up to the threshold weight or decreasing them from a higher weight through some form of leniency or plea bargaining mechanisms. Bjerk (2017b) suggests that bunching likely occurs from downward manipulation while Tuttle (2025) 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. Receiving a conspiracy charge often connects a defendant to a larger quantity of drugs or provides more evidence of high drug quantity involvement. Conspiracy is also 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.

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.

This may show up by way of bunching or by way of MM charge application.

Tuttle (2025) gives evidence of the significance of conspiracy charges in case manipulation. 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. I find that for cases with weights 50% to 99% of the threshold weight, about 44% carry a conspiracy charge. Likewise, for cases with weights just above the threshold (101% to 150% of the threshold), 48% receive a conspiracy charge. But at the bunching point, the percent of cases carrying conspiracy charges is 83%. This is illustrated in Figure A.2a in the appendix.

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** - sometimes also called the conspiracy-wide approach, this 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 or MM charging dependant upon which circuit they practice 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.[4] 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.[5]

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 2 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,[6] and the latest being in June of 2018.[7] 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.

II. C.   Policy Details

Here I provide a brief description of the three main policies used, detailing what changes they brought to charging or sentencing in the context of mandatory minimums.
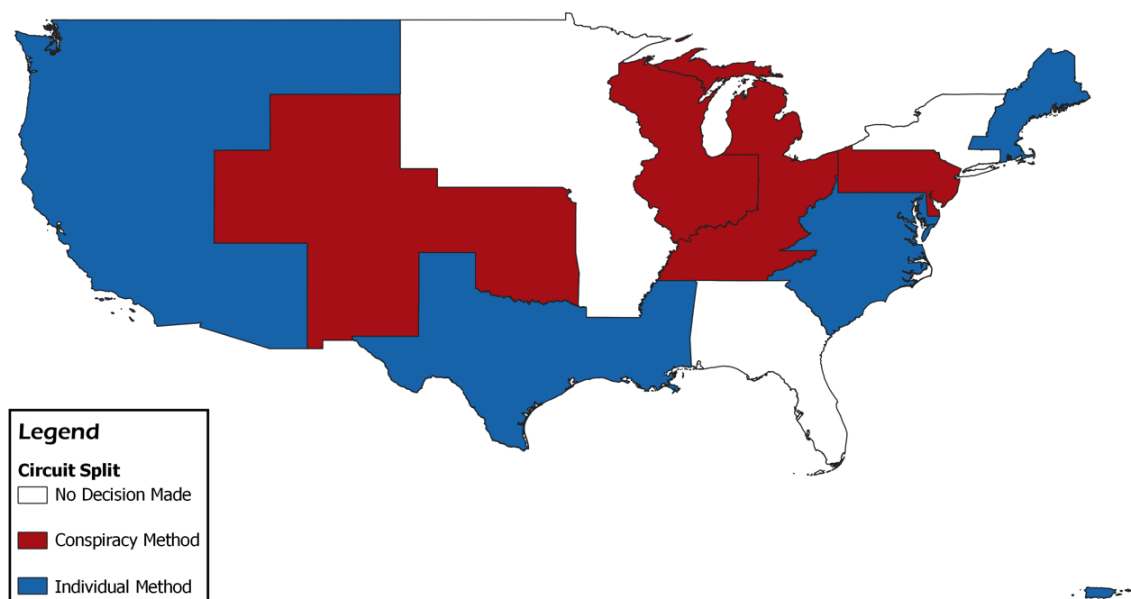
---

[4]Conspiracy clearly still has an effect on manipulation; when comparing conspiracy charges by weight, the strong bunching seen in the appendix Figure A.2a is present across each regime.

[5]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).

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

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

Figure 2: Conspiracy Drug Weight Approach by Circuit

The Fair Sentencing Act

The Fair Sentencing Act of 2010 (P.L. 111–220) was signed into law in August of 2010. This law increased the mandatory minimum required threshold weight for crack (base) cocaine. For 10-year MM charges, the threshold weight was raised from 50 grams to 280 grams. This means many cases would no longer be eligible for MM charges under the new weight requirements.

This law was passed to explicitly combat drug and race disparities in MM charges. Prior to Fair-2010, the disparity in threshold weights between powder and crack cocaine was 100:1. Fair-2010 reduced this to 18:1. Because the vast majority of crack cases are from Black defendants and crack cases often carry high sentence lengths, the law was expected to help reduce race gaps in charging and sentencing.

The Drugs Minus 2 Amendment

The Drugs Minus 2 Amendment was a major decrease in punishment levels for most drug trafficking cases. When determining charges, baseline offense levels for specific quantities of drugs are prescribed from the drug quantity table. Minus 2 decreased these offense levels by two for quantity ranges that included the MM threshold weights. All other weight ranges were then extrapolated from these points. This essentially lowered the expected sentence

length for all trafficking cases with low enough criminal history levels. Defendants considered "career offenders" were excluded from the Minus 2 amendment as they are sentenced outside of the guidelines of the quantity table. The United States Sentencing Commission (USSC) predicted Minus 2 would impact around 70% of all trafficking cases.

Because the offense levels provide windows of sentence lengths, the decrease in sentence lengths caused by Minus 2 is difficult to predict, especially if prosecutors or judges responded to the amendment with harsher charging or sentencing. However, a reduction of two levels may induce a large decrease in sentence length. The USSC predicted a sentence length decrease of 17.7% from pre-Minus 2 means.

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.[8]

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 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

---

[8]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.

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   Data and Empirical Strategy

III. A.   Data

The primary drug case data is provided by the United States Sentencing Commission (USSC) and includes all federal drug trafficking cases from 2008 to 2020. Data is restricted to the five most prevalent substances subject to mandatory minimum sentencing: powder cocaine, crack, heroin, methamphetamine, and marijuana.[9] I also restrict the data to the primary racial groups of study: White, Black, and Hispanic. Because prosecutor manipulation and application decisions are most likely to impact cases close to the threshold weight, I limit the main analysis to 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, though wider ranges of weights are considered in the robustness checks section. 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 27,832 observations over the 12 year period. However, each regression analysis considers only four years at a time; two years pre-policy and two years post. These 4-year spans are necessary to avoid policy timing overlap in pre-periods or post-periods.

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

---

[9]The other three substances subject to mandatory minimum law are PCP, LSD, and fentanyl. These had too few observations during the analyzed time periods for any meaningful analysis.

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 about half of all cases in the sample, while conspiracy-circuits make up just under 25% of all cases. Panel A. gives the distribution of drug type charges. Drug types are represented fairly equally with the most common drug type being cocaine, making up 26.5 percent of cases. Panel B. gives statistics for defendant characteristics. Over 78 percent of the sample is made up of minority-race defendants and 86.8 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 80 months, signifying the high punishment level for these types of drug crimes. Across each panel, figures are very close between circuit regimes, indicating differences in results are not driven by major differences in case composition.

In addition to the USSC data, I use census data to control for racial population changes over time at the district 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.[10] I also use google trends search data from Stephens-Davidowitz (2014) to consider how my results correlate geographically with racial animus.

III. B.   Empirical Strategy

Each of the three policies studied impact prosecutor incentives, either by reducing the expected impact of MM eligibility, or by increasing the cost to achieving MM eligibility. To compare these effects across circuit type and race, I employ a difference-in-differences approach utilizing the timing of each policy. The following model captures the differential

---

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

impact of each policy on racial minorities compared to White counterparts:

$$Y_{idtm} = \alpha + \beta_1 minority_i + \beta_2 post_{tm} + \beta_3 minority_i \cdot post_{tm} + X_i\gamma + \lambda_d + \delta_m + \epsilon_{idtm} \quad (1)$$

The model is run individually for conspiracy-method circuits and for individual-method circuits to test policy impacts on race disparities across evidentiary regimes. Here Y_{idtm} measures the outcome of interest for case i in district d during year t and month m. The primary tested outcomes are a binary measure for whether case i was sentenced at the threshold weight and a binary measure for whether case i had the MM application. X_i is a vector of defendant and case characteristics, including the drug type, criminal history points, age, age squared, a binary measure for whether the defendant has a college education, and a binary measure for whether the defendant is an illegal alien. District and month-of-year fixed effects are included. I omit monthly fixed effects in the main difference-in-difference models so that the full impacts of the policy can be observed in the regression output. 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.

To test whether changes in race disparities are statistically different across circuit groups, I employ a triple-differences model as follows:

$$Y_{idtm} = \alpha + \beta_1 minority_i + \beta_2 post_{tm} + \beta_3 conspiracy\_circ_d + \beta_4 minority_i \times post_{tm}$$
$$+ \beta_5 minority_i \times conspiracy\_circ_d + \beta_6 post_{tm} \times conspiracy\_circ_d + \beta_7 ddd_{idtm}$$
$$+ X_i\gamma + P_{dt}\omega + \lambda_d + \tau_{tm} + \epsilon_{idtm} \quad (2)$$

This model now includes monthly fixed effects to control for district invariant factors that may impact results. I also include a vector P_{dt} which includes annual racial population percentages by district. Including these controls ensures results are driven by changes in the racial makeup of districts. β_7 provides the desired effect - the differential impact of the policy on race gaps across circuit groups.

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 wider ranges of cutoffs 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, as individuals who have 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. Regarding departure, I find that minority defendants do receive less government sponsored sentence reduction than Whites defendants, implying a lower willingness to cooperate. However, this difference in departure rates is largest in the early years of the sample and for non-crack cases. By year 2012, average government departure rates for Whites is around 43% of cases and around 36% for minorities. Finally, 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.
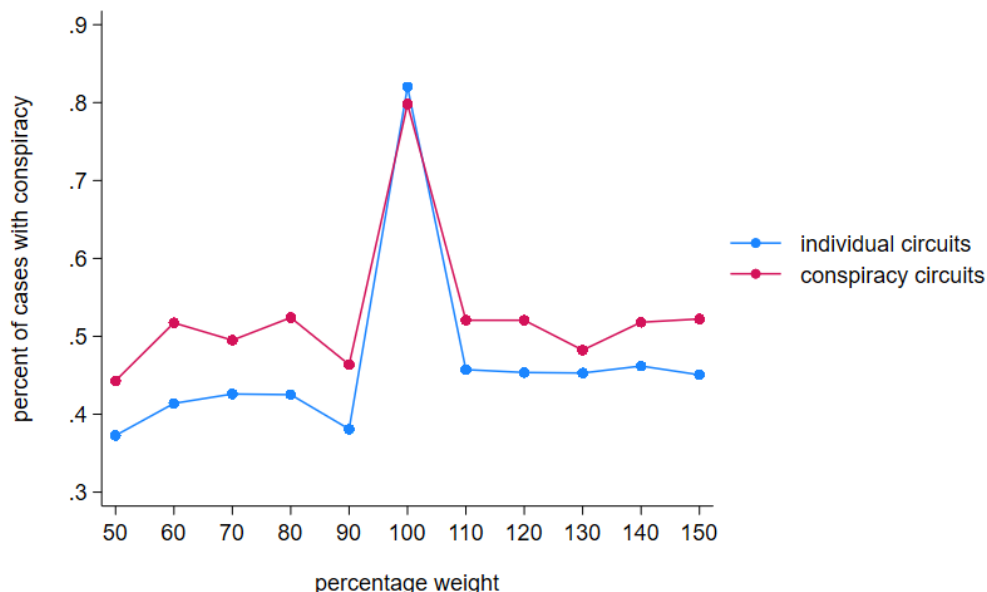
IV Results

I first present some descriptive statistics which consider the overall differences in MM outcomes between the two circuit regimes. These descriptives illustrate whether circuit regimes have an observable effect MM sentencing and prosecutor decisions to seek higher eligibility. Following these stylized facts I present the main difference-in-difference and triple difference results to show the effect of each policy on the race disparities between each regime.

IV. A. Descriptive Statistics

If conspiracy-method evidentiary rules incentivize prosecutors to pursue MM eligibility more frequently, we would expect higher levels of conspiracy charges in conspiracy circuits. In Figure A.1, I compare the percent of cases with conspiracy charges at weights local to the 10-year MM eligibility cutoff. I find that conspiracy charge rates are higher at all weights except at the bunching point. The spike in conspiracy rate at this point indicates that prosecutors in both circuit types engage in drug weight manipulation activity and that conspiracy charges are the primary mechanism for manipulation under each regime. Higher conspiracy rates at weights under the threshold may indicate more attempts at drug weight manipulation while weights above may indicate weights that are pushed above the threshold weight. In the appendix, I show that this gap in conspiracy rates is present across the entire timeline and that it is driven by both races, but with a larger gap in White cases pre-FSA and the gap being entirely driven by minority cases post-FSA. This is shown in Figure A.2b.

Equal conspiracy rates at the bunching point is not particularly surprising; these cases reflect instances where the prosecutor was able to push the weight up successfully. However, if manipulation is easier in conspiracy circuits, we would expect a higher proportion of cases to be pushed to this bunching point in the conspiracy circuits compared to the individual circuits. Indeed, I find that across the entire timeline, bunching percentages are higher in conspiracy circuits, particularly for years following the Minus-2 Amendment. This is illustrated in Figure 4. This gap is most prominent between the timing of the Minus-2 Amendment and the FSA. In the appendix, I consider the bunching patterns in this time period separately by race. This is seen in Figure A.3. Contrary to the conspiracy results,
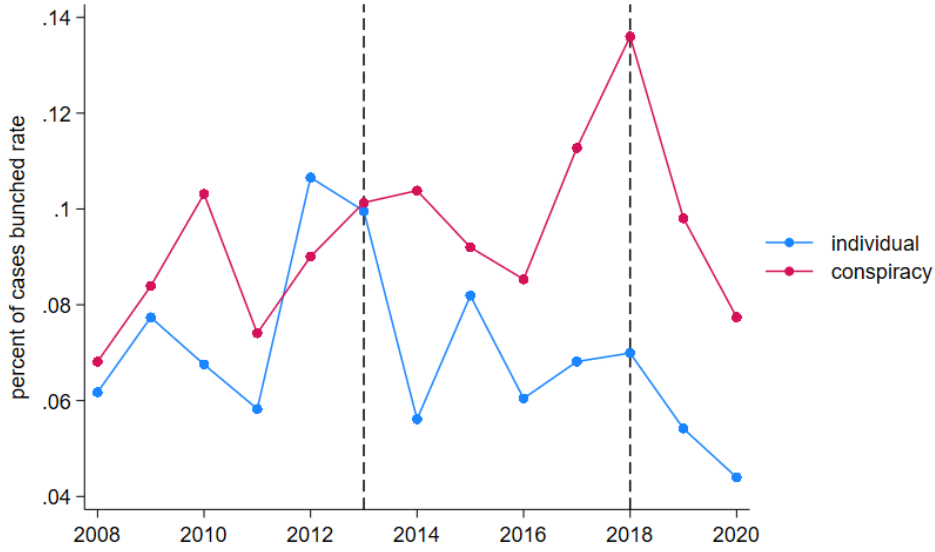
Figure 3: Conspiracy Charges by Weight

Notes: 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.

the gap between conspiracy and individual circuits is driven primarily by minority cases. I likewise check overall MM application rates in each circuit and find rates are slightly higher in conspiracy circuits.

The higher rates of conspiracy charges, eligibility bunching, and MM application in the conspiracy circuits are evidence that the evidentiary rule difference created by the circuit split has an effect on prosecutor charging decisions. That is, when costs to additional evidence gathering are lower, prosecutors do indeed pursue higher charges through conspiracy charges that increase drug weights. The descriptive statistics suggest this is pursued more often for White defendants as conspiracy charge disparities are driven by White cases, but the actual outcomes may impact minorities more, who drive the gap in bunching. To more effectively tease out race disparity effects, I now turn to the policy analysis.

Figure 4: Overall Bunching by Circuit Type

IV. B.   Race Disparity Analysis

Fair Sentencing Act Results

Because Fair-2010 specifically impacted crack-cocaine cases, the following analysis only concerns cases where crack is the drug type nearest to the MM threshold or was the primary drug type. Furthermore, all results focus on the new cutoff of 280 grams. I find that Fair-2010 effectively reduced race disparities in formal MM charges in conspiracy method circuits, while it slightly widened the race gap for circuits in the individual regime. Before the policy, formal MM charges in individual circuits occurred at about the same rate for White and minority defendants. This changes post policy with an increase in race disparities, though these results are noisy. Conversely, conspiracy circuits exhibit large race gaps prior to the policy which are effectively reduced by Fair-2010. The difference-in-difference estimates for these results are presented in Table 1. Note that changes in bunching follow similar directions, but with smaller magnitudes and without statistical significance. Differences in these race-gap results are significant across districts as shown by the triple-differences results.

The existing race gap in conspiracy circuits was closed by Fair-2010, but a race gap was created in individual circuits as benefits of the policy disproportionately went to White defendants. This is a pattern that will continue to manifest in the other two policy analyses

17

- that when a race gap is present, the policy will disproportionately benefit minorities in a way that eliminates the race disparity. But when a there is no pre-policy race gap, benefits will go disproportionately to White defendants.

Finally, these results measure qualitative differences in crack cases pre and post-Fair-2010, but do not account for large changes in the number of cases. Indeed, Fair-2010 led to a significant decrease in the number of crack cases in both conspiracy and non-conspiracy regime circuits. Appendix Figure A.4 shows the percentage of all drug cases with crack as the primary drug type. The percent of crack cases decreases from a pre-Fair-2010 level of about 30% for non-conspiracy circuits and 40% for conspiracy circuits down to about 13% for each regime by the year 2014. So while conspiracy circuits pursued more crack cases overall, this gap was also closed by Fair-2010 after prosecutors were significantly less likely to pursue crack cases.

Drugs Minus 2 Amendment Results

Table 2 displays the results for the Minus-2 amendment analysis. Minus-2 had a similar impact on race disparities as Fair-2010; larger race gaps in conspiracy circuits were reduced by the amendment while race gaps increased slightly in non-conspiracy circuits. Table 2 reports the difference-in-difference estimates. In non-conspiracy circuits, pre-period race gaps were effectively nonexistent and Minus-2 led to significant decreases in both bunching and MM charges, with a slightly larger decrease for White defendants. In conspiracy circuits however, minority defendants were nearly 10 percentage points more likely to receive a MM charge compared to their White counterparts before Minus-2. This gap was entirely closed following Minus-2 as minorities experienced drastic decreases in MM charge rates while White defendants were unaffected. This pattern is illustrated in Figure 5 with the corresponding event studies shown in Appendix Figure A.5. Bunching effects follow a similar pattern but do so with minimal statistical significance.

Table 1: Fair-2010 Analysis

Panel A: DiD	**Individual Circuits**		**Conspiracy Circuits**	
	(1)	(2)	(3)	(4)
	bunch	MM app.	bunch	MM app.
minority	-0.115	-0.0673	-0.00568	0.0953
	(0.104)	(0.0844)	(0.00463)	(0.0574)
post	-0.0468	-0.318**	0.00164	0.146
	(0.132)	(0.125)	(0.00846)	(0.139)
post × minority	0.0864	0.137	0.0179	-0.277**
	(0.126)	(0.128)	(0.0122)	(0.126)
baseline controls	yes	yes	yes	yes
monthly fixed effects	no	no	no	no
population controls	no	no	no	no
N	936	936	751	751

Panel B: Triple-Diff	(1)	(2)	(3)	(4)
	bunch	MM app.	bunch	MM app.
post × minority × consp_circ	-0.0402	-0.554***	-0.0545	-0.309*
	(0.0689)	(0.176)	(0.0774)	(0.152)
baseline controls	yes	yes	yes	yes
monthly fixed effects	no	no	yes	yes
population controls	no	no	yes	yes
N	1687	1687	1687	1687

Notes: Panel A reports the difference-in-difference estimates for conspiracy and individual circuits separately, while Panel B reports the triple-differences coefficients. The dependent variable "bunch" is a binary variable that equals one if the charging weight is exactly equal to the mandatory minimum threshold weight. "MM app." is a binary variable that equals one if the case carries the MM application. Baseline controls include primary drug type, defendant sex, criminal history points, age, age squared, education, citizenship status, district fixed effects, and month-of-year fixed effects. Population controls include racial population percentages at the district-year level. Standard errors are clustered at the district level.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Minus-2 Analysis

Panel A: DiD	Individual Circuits		Conspiracy Circuits	
	(1)	(2)	(3)	(4)
	bunch	MM app.	bunch	MM app.
minority	-0.00334	-0.0193	0.00989	0.0957**
	(0.0137)	(0.0241)	(0.0240)	(0.0397)
post	-0.0379**	-0.133***	0.0290	-0.0154
	(0.0177)	(0.0357)	(0.0262)	(0.0310)
post × minority	0.0182	0.0324	-0.0485	-0.110**
	(0.0217)	(0.0384)	(0.0368)	(0.0469)
baseline controls	yes	yes	yes	yes
population controls	no	no	no	no
monthly fixed effects	no	no	no	no
N	4378	4378	2115	2115

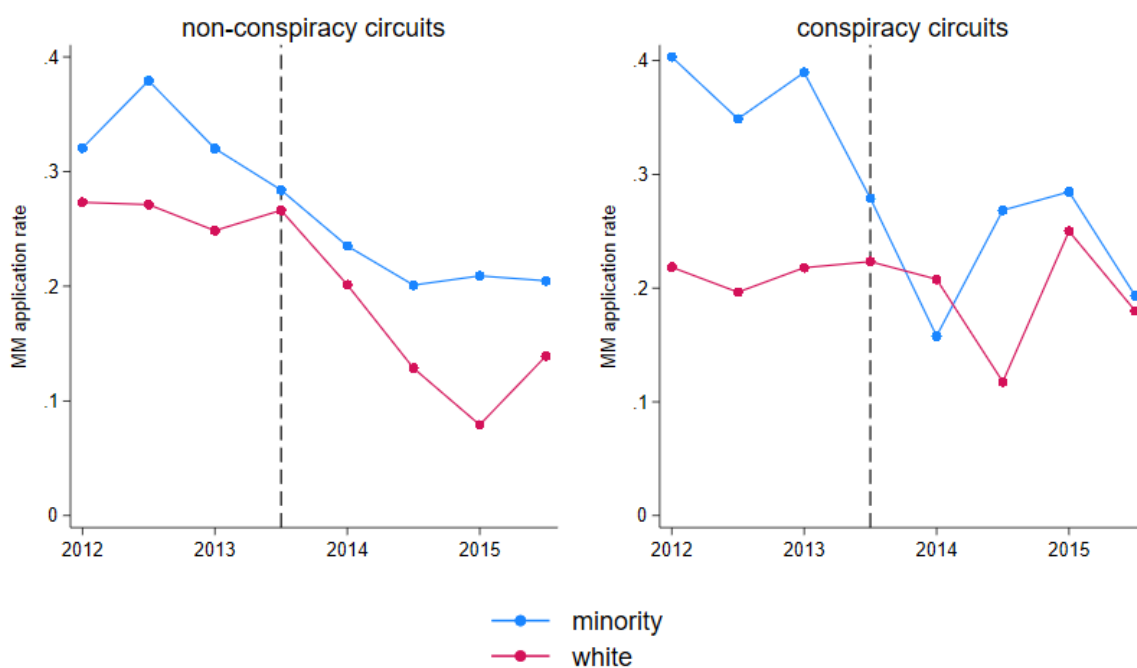| *Panel B: Triple-Diff*	(1)	(2)	(3)	(4)
	bunch	MM app.	bunch	MM app.
post × minority × consp_circ	-0.0683	-0.125**	-0.0676	-0.114*
	(0.0466)	(0.0612)	(0.0474)	(0.0579)
baseline controls	yes	yes	yes	yes
population controls	no	no	yes	yes
monthly fixed effects	no	no	yes	yes
N	6493	6493	6493	6493

Notes: Panel A reports the difference-in-difference estimates for conspiracy and non-conspiracy circuits separately, while Panel B reports the triple-differences coefficients. The dependent variable "bunch" is a binary variable that equals one if the charging weight is exactly equal to the mandatory minimum threshold weight. "MM app." is a binary variable that equals one if the case carries the MM application. Baseline controls include primary drug type, defendant sex, criminal history points, age, age squared, education, citizenship status, district fixed effects, and month-of-year fixed effects. Population controls include racial population percentages at the district-year level. Standard errors are clustered at the district level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

As described in Section II. C., the Drugs Minus 2 Amendment occurred at a similar timing as two other events that changed MM sentencing: the *Alleyne v. United States* decision and the Holder Memo. These two events occurred before Minus-2 and may also contribute to the effects reported above, including apparent anticipatory effects seen in Figure 5. Because the timing of these events is close, it is difficult to disentangle one from another. However, in an attempt to do so, I run the regression analysis again using the timings of *Alleyne* and

Holder, reported in appendix Table A.2. Note that for the Holder regressions, only cases eligible under the memo description are included. Under these two alternate timings, the non-conspiracy circuit results disappear with no real race gap changes and no significant effects for White or minority defendants. Each regression for conspiracy circuits reports a similar story to the main results, with substantial decreases in race gaps, though these results are only significant at the 10% level. Regardless of which policy drives the effect, these collection of events mitigated large existing race disparities in conspiracy circuits and marginally increased race disparities in non-conspiracy circuits.

Figure 5: Minus-2 event study for MM application

First Step Act Results

As mentioned in the descriptive statistics section, pre-FSA bunching exhibited large race disparities in the individual circuits with cases at the bunching point being almost entirely made up of minority defendants. In the conspiracy circuits, bunching was nearly equal between White and minority defendants. The FSA caused a large, immediate drop in bunching at the threshold weight across both circuits, with relatively little impact on MM
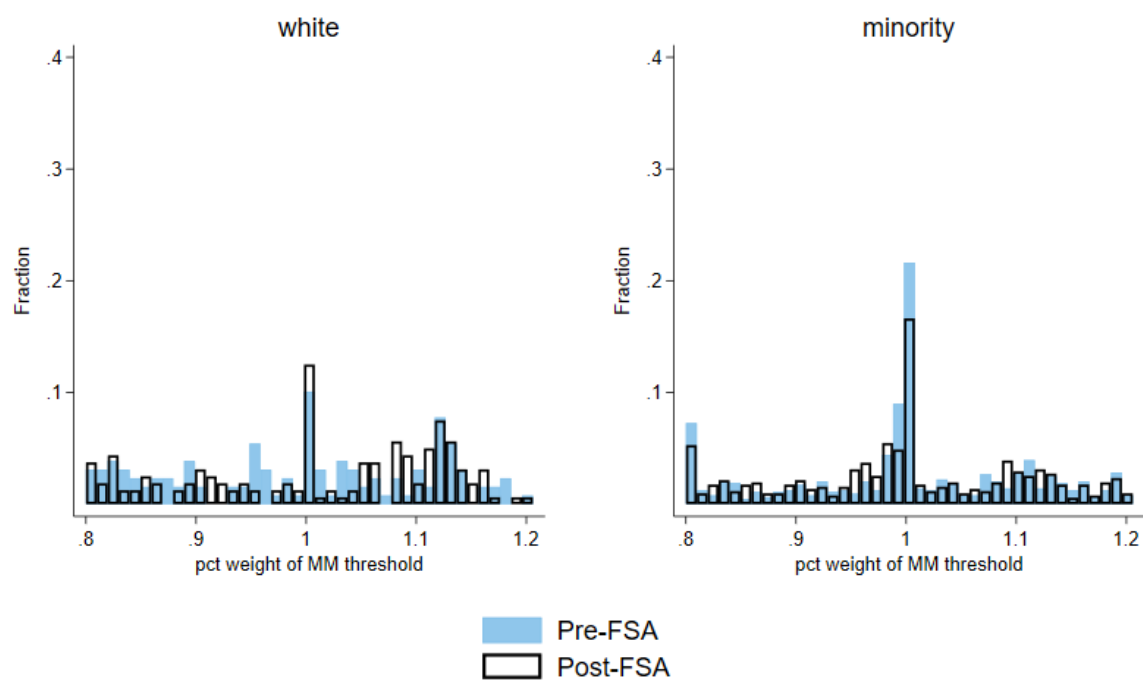
application across either circuit or race group.

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 3 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. The regression results corroborate the effects illustrated in Figure 6. For minority cases in the individual circuits, bunching decreased around 3.08 percentage points, with White cases experiencing a small, insignificant increase in cases at the bunching point. In conspiracy circuits the FSA significantly reduced bunching for White cases, with a decrease of 9.76 percentage points. The difference-in-differences estimate shows a difference in effects between White and minorities of 10.1 percentage points, indicating no real effect on minorities. Event studies graphs for FSA analysis are given in appendix Figure A.6.

Figure 6: Bunching effects of FSA by circuit group and race
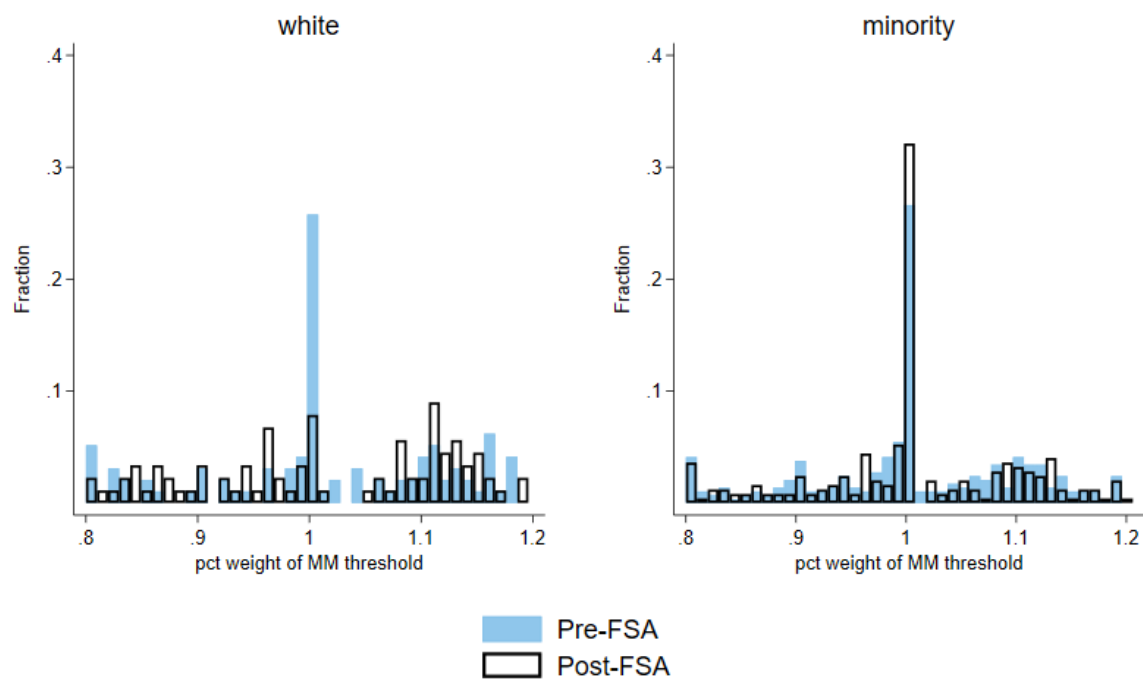
(a) Individual Circuits

(b) Conspiracy Circuits

Table 3: FSA Analysis

Panel A: DiD	Individual Circuits		Conspiracy Circuits	
	(1)	(2)	(3)	(4)
	bunch	MM app.	bunch	MM app.
minority	0.0194	0.00323	-0.0563*	0.00302
	(0.0140)	(0.0235)	(0.0327)	(0.0331)
post	0.0127	0.0482*	-0.0976***	0.0362
	(0.0117)	(0.0248)	(0.0310)	(0.0483)
post × minority	-0.0308*	0.00129	0.101**	-0.00785
	(0.0155)	(0.0206)	(0.0397)	(0.0519)
baseline controls	yes	yes	yes	yes
population controls	no	no	no	no
monthly fixed effects	no	no	no	no
N	3199	3199	1519	1519

Panel B: Triple-Diff	(1)	(2)	(3)	(4)
	bunch	MM app.	bunch	MM app.
post × minority × consp_circ	0.141***	-0.0205	0.134***	-0.0145
	(3.42)	(-0.40)	(3.08)	(-0.24)
population controls	yes	yes	yes	yes
baseline controls	yes	yes	yes	yes
monthly fixed effects	no	no	yes	yes
N	4718	4718	4718	4718

Notes: Panel A reports the difference-in-difference estimates for conspiracy and non-conspiracy circuits separately, while Panel B reports the triple-differences coefficients. The dependant variable "bunch" is a binary variable that equals one if the charging weight is exactly equal to the mandatory minimum threshold weight. "MM app." is a binary variable that equals one if the case carries the MM application. Baseline controls include primary drug type, defendant sex, criminal history points, age, age squared, education, citizenship status, district fixed effects, and month-of-year fixed effects. Population controls include racial population percentages at the district-year level. Standard errors are clustered at the district level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Taken together, the results show that (1) the levels of these charging measures are higher in conspiracy charges and that after each policy, (2) when race disparities are present pre-policy, policies reduce race gaps as benefits mostly go to minority cases, and (3) when race disparities are not present pre-policy, policies create race gaps as benefits mostly go towards White cases. One way to think of this is that when drug policy is passed, the benefits first go to White defendants, regardless of circuit regime, and that conspiracy circuits "lag behind"

the individual circuits. The fact that the marginal cases group by race suggests one of three potential mechanisms. First, minority cases may have lower costs to gathering evidence. This could occur by way of differences in criminal activity, gang affiliation, and quality of legal counsel. Second, the benefits to gathering additional evidence may be higher for minorities than for Whites. This would occur if additional evidence leads to higher average sentence lengths for minorities compared to Whites. This mechanism is supported by the companion paper to this piece, Cooper (2025). Finally, the results could be explained by taste-based discrimination, where prosecutors decrease fact-finding for White cases more than minorities simply due to preference. Regardless of the mechanism, this paper shows that marginal cases tend to be grouped by race with White individuals being the first to benefit.

Direction of Manipulation

While the strong correlation between bunching and conspiracy charges indicates that manipulation is most likely occurring upward (where prosecutors push cases from below the threshold weight up to it), I also visually check whether the policy results support this assumption. This is done by considering the changes in the distribution following the FSA for the groups that experience the largest decreases in bunching - minority cases in conspiracy circuits and White cases in individual circuits. Specifically, I look for whether the distribution increases on the left or right of the threshold weight, with the assumption that before the policy these cases would have been bunched. If prosecutors are indeed engaging in fact-finding in a way that increases drug weights, the expected effect is an increase mass to the left of the distribution.

Figure A.9 presents a kernel density graph before and after the FSA for the marginal groups. The figure shows the excess mass appearing on the left side of the distribution very close to the threshold weight, suggesting bunching is a product of cases being manipulated upwards. Additionally, the graph suggests that the cases that are manipulated are those with weights already very close to the MM threshold. This is intuitive as these cases are likely to have the lowest cost to manipulation as less additional drug weight is needed to ensure MM eligibility.

IV. C.   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 the use of wild-cluster bootstrapped standard errors, the inclusion of circuit-specific time trends, expanding the individual circuits to include all non-conspiracy circuits, whether results are driven by cartel-heavy geographies, and whether results hold when using larger samples that include a wider range of drug weights. These results are presented in appendix Tables A.3 and A.4. To reduce the number of tables I do each of these analyses using the triple-differences design, each of which includes the full set of controls.

Throughout the analysis, I cluster on the district. Wild-cluster bootstrapped standard errors may be necessary due to relatively low number of clusters - 29 in conspiracy circuits and 36 in individual circuits. Panel A of Table A.4 presents the triple-difference estimates with wild-bootstrapped standard errors. The main results significance levels remain unchanged. Panel B includes the specifications with race-specific time trends included (minority or white), while Panel C gives results with circuit-specific time trends included. In either case, results remain statistically significant and similar in magnitude to the main results. Finally, Panel D runs the analysis again while dropping circuits that have especially high drug cartel presence according to the 2019 DEA National Drug Threat Assessment. It's worth noting that none of these districts are all in conspiracy-method circuits. The triple-difference estimates remain significant for the FSA results, while the Minus-2 and Fair-2010 results lose significance, but are similar in magnitude.

I also consider whether the results change when I include circuits for which no explicit ruling has been made regarding drug weights and conspiracy charges. I assume these circuits likely follow the individual method as *Alleyne v. United States* requires all facts increasing a mandatory minimum to be determined by a jury, a higher burden of proof.[11] The descriptive statistics show less of a disparity when including these circuits, but the race gap results actually get stronger. The difference in race-circuit outcomes gains stronger magnitudes and

---

[11]This does not necessarily imply the individual method, which is why I omit these circuits from the main analysis. One may also wonder whether *Alleyne* might have effectively reversed the conspiracy-method rulings. Several circuits have had additional challenges following *Alleyne* under which the court decided not to reverse previous conspiracy-method rulings.

become more statistically significant.[12]

Finally, because the 50%-150% window is arbitrary, I check whether results are not driven by biased sampling. To do this, I run the triple-differences model for each policy again using a two different samples; one that includes 20%-200% of the threshold weight and one that includes all cases below 300% of the threshold weight. Note that the 0%-300% window includes the lower MM threshold as well, as well as many cases with drug weights close to 0. While the magnitudes of these findings shrink, the significance levels are nearly identical, with only the Minus-2 MM application result falling to 10% significant in the largest sample. Overall, the main findings are consistent through each of the robustness checks provided.

V   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 show that circuits with lower costs to drug weight evidence gathering exhibit higher levels of conspiracy charges, a higher proportion of cases at the bunching point, and a higher rate of MM application. This indicates that prosecutors are responsive to the lower costs created by the circuit split, and that their pursuit of higher charges results in stronger punishments at sentencing. I then show how racial groups are differentially impacted by three major MM policies across the circuit regimes. Across each of the three policies a race ratcheting effect is present where benefits of the policy are driven by one racial group and White defendants receive benefits first. Conspiracy circuits lag behind individual circuits in

---

[12]Once again, these results are omitted for space but are available upon request.

this pattern.

On a broader scale this paper suggests that prosecutors are sensitive to changes regarding this evidence-sentence tradeoff. If reducing existing racial disparities is a goal for policymakers, the policy recommendation from this paper is to increase evidence costs or reduce sentence length impacts. However, policymakers should be aware that such changes may unintentionally create race disparities in geographies where race gaps are not present, since benefits will disproportionately favor White defendants. 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. This behavior may be driven by differences in costs or benefits in seeking evidence between each racial group.

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.

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. (2023). *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.

Cooper, S. (2025). *How much do mandatory minimums matter?* (Working Paper). University of Connecticut.

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. (2024). *Charging leniency 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*.

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.

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

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

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.

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.

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. (2025). *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.

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.[13] 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.[14] 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

---

[13]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.
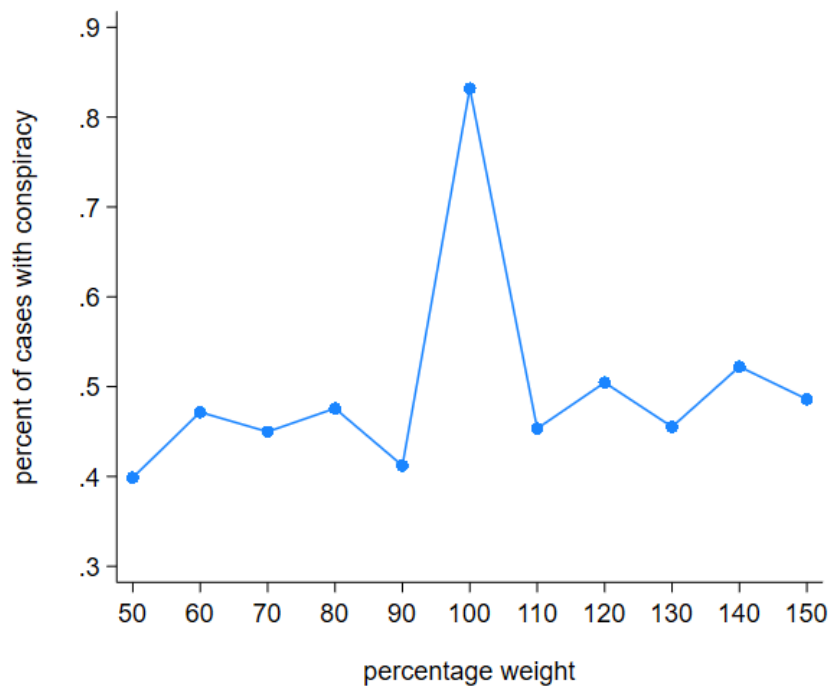
[14]The details of which crimes qualify as "serious" are more technical than explained here, but include most drug or violent crimes for which the offenders was imprisoned for over 12 months. See Public Law 115-391 for more details of seriousness as well as which statutes qualify as a violent crime.

weapon in connection with the crime, the offense did not result in death or serious injury 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.[15] 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.

---

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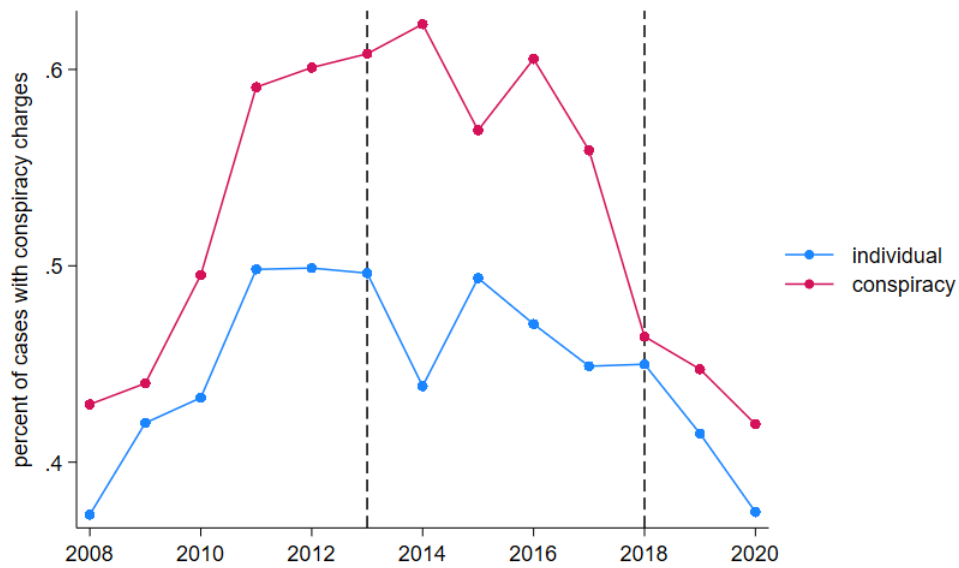
Additional Figures

Figure A.1: 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.

Figure A.2: Percent of cases with a conspiracy charge by year
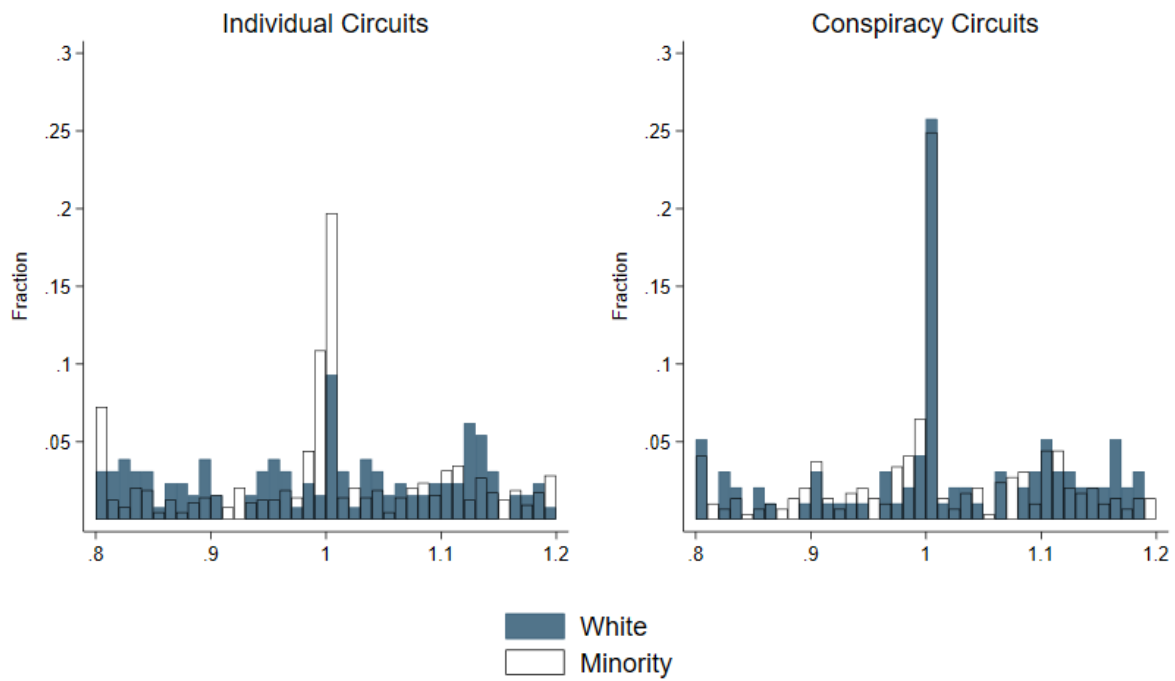
(a) All cases

(b) By race

Figure A.3: Percent of cases bunched by race

The following histogram present distributions by race-circuit-type intersection for two years preceding the timing of the First Step Act.
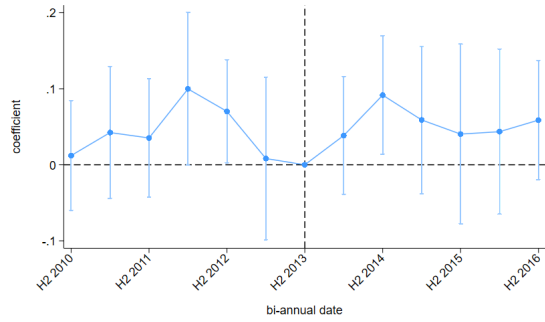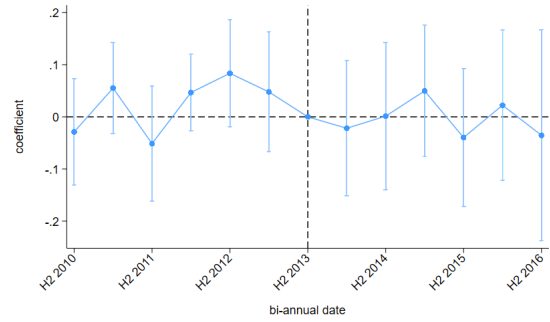
Figure A.4: Proportion of cases crack

This graph shows what proportion of drug cases had crack cocaine as the primary drug type by year and circuit type.
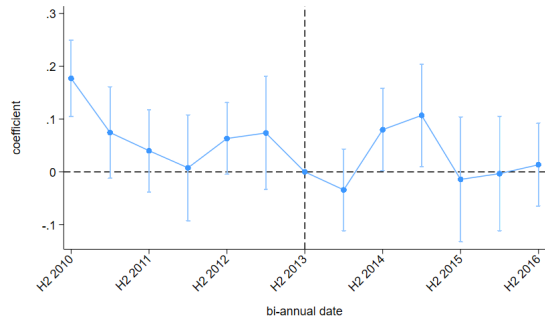
Figure A.5: Minus-2 event studies

(a) Bunching, individual circuits

(b) Bunching, conspiracy circuits

(c) Charges, individual circuits
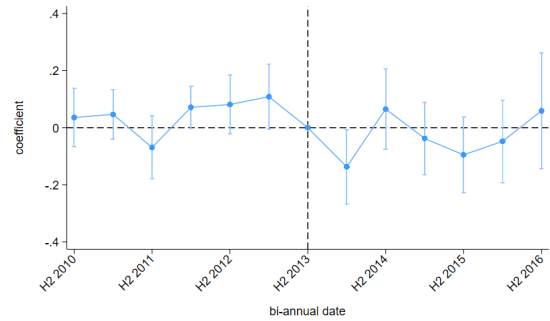

(d) Charges, conspiracy circuits

Figure A.6: FSA event studies

(a) Bunching, individual circuits

(b) Bunching, conspiracy circuits

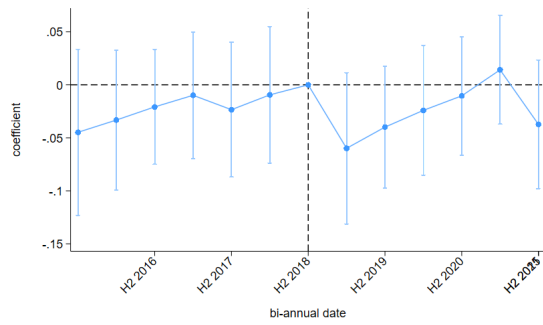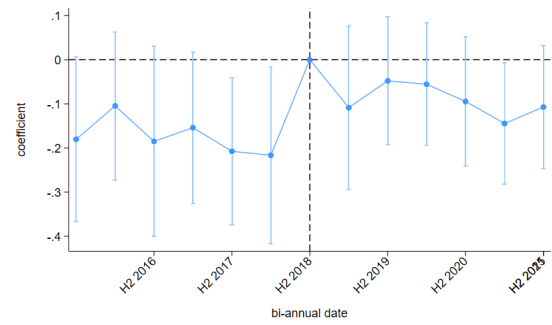**(c)** Charges, individual circuits

(d) Charges, conspiracy circuits

Figure A.7: FSA kernel density
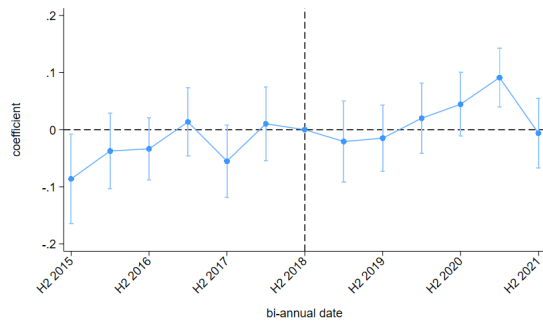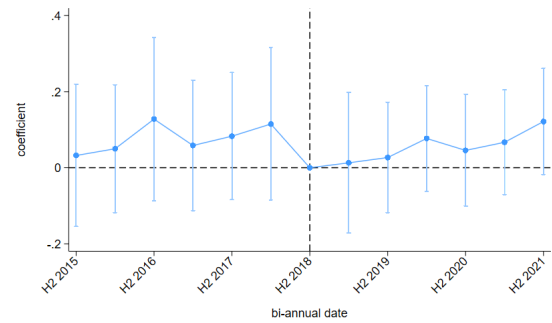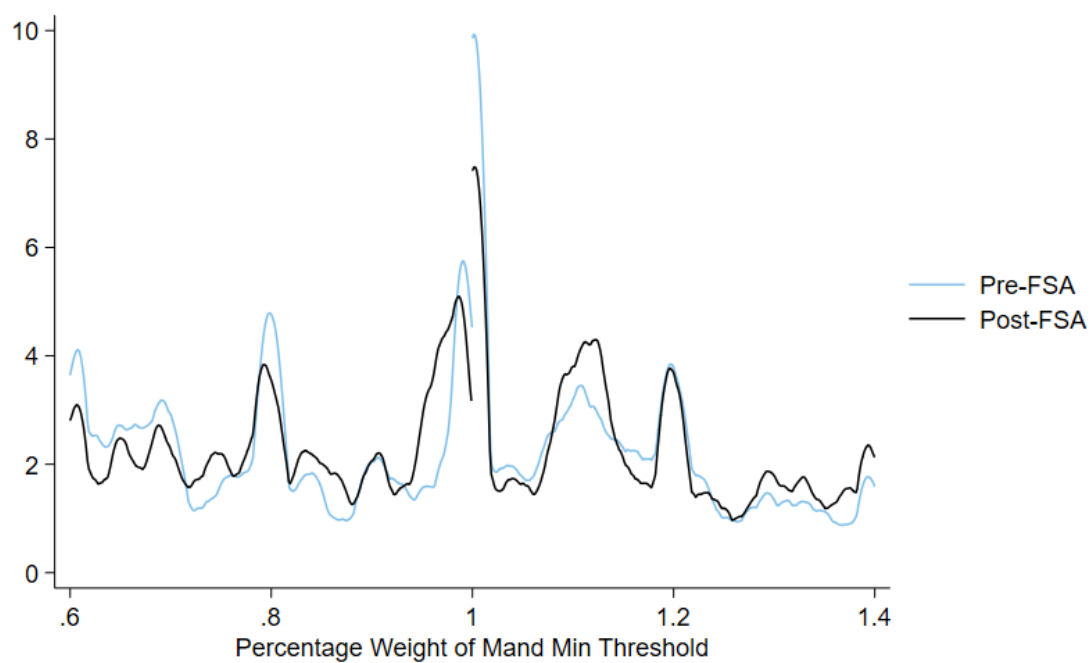

Notes: This figure combines White-conspiracy cases with minority-individual cases to see where excess mass arises following the FSA. These two groups are combined as they are the cases that experience the largest decrease in bunching throughout the analysis.

Additional Tables

Table A.1: USSC Drug Case Data - Summary Statistics

	All Cases	Individual Circuits	Conspiracy Circuits
Panel A. Drug Types			
cocaine	0.265	0.269	0.245
crack	0.206	0.171	0.250
heroin	0.144	0.122	0.169
marijuana	0.136	0.201	0.0856
meth	0.250	0.237	0.250
Panel B. Defendant Characteristics			
minority	0.780	0.810	0.748
female	0.132	0.140	0.119
age	35.17	34.81	35.22
criminal history points	3.943	3.654	4.422
college	0.183	0.169	0.181
illegal alien	0.161	0.188	0.141
Panel C. Outcomes			
percent weight of threshold	0.920	0.919	0.920
bunched (binary)	0.0914	0.0713	0.0923
10-year MM applied	0.282	0.266	0.298
conspiracy charge	0.490	0.449	0.526
safety valve applied	0.320	0.348	0.265
assistance to government	0.283	0.297	0.297
sentence length (in months)	80.36	78.43	84.19
Share	1	0.500	0.247
Observations	27832	13924	6872

Table A.2: Alternative timing results for Minus-2

	Individual Circuits		Conspiracy Circuits	
	(1)	(2)	(3)	(4)
	bunch	MM app.	bunch	MM app.
Panel A: Alleyne Results				
minority	0.0128	0.0122	0.00133	0.0846*
	(0.0126)	(0.0430)	(0.0224)	(0.0492)
post	-0.0147	-0.0734	0.0410*	-0.0619
	(0.0208)	(0.0602)	(0.0210)	(0.0418)
post × minority	-0.0022	-0.0249	-0.0365	-0.0999*
	(0.0219)	(0.0536)	(0.0273)	(0.0536)
N	4790	4790	2288	2288
Panel B: Holder Results				
minority	-0.0119	-0.03046	0.0251	0.0459
	(0.0245)	(0.0334)	(0.0254)	(0.0526)
post	-0.0288	-0.0827	0.0793**	-0.0502
	(0.0285)	(0.0689)	(0.0382)	(0.0687)
post × minority	0.0123	-0.0215	-0.0955*	-0.138*
	(0.0284)	(0.0532)	(0.0514)	(0.0767)
N	2624	2624	1090	1090

Notes: This table is analogous to Table 2 in the main text, but uses the timing of *Alleyne* and Holder instead, with results for *Alleyne* presented in Panel A and Holder in Panel B. Note that the Holder analysis only contains cases that are eligible under the Holder decision. Each regression controls for primary drug type, defendant sex, criminal history points, age, age squared, education, citizenship status, and district and month of year fixed effects. Standard errors are clustered at the district level.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Robustness - triple differences

	Fair-2010		Minus-2		FSA	
	(1)	(2)	(3)	(4)	(5)	(6)
	bunch	MM app.	bunch	MM app.	bunch	MM app.
Panel A: *Wildbootstrap Clusters*						
ddd	-0.0545	-0.309*	-0.0676	-0.114*	0.134***	-0.0145
	(0.128)	(0.173)	(0.0476)	(0.0582)	(0.0439)	(0.0599)
Panel B: *Minority-Specific Time Trend*						
ddd	-0.0504	-0.312*	-0.0707	-0.121**	0.133***	-0.0139
	(0.121)	(0.165)	(0.0472)	(0.0574)	(0.0434)	(0.0597)
Panel C: *Circuit-Specific Time Trends*						
ddd	-0.0673	-0.310*	-0.0755	-0.133***	0.138***	-0.0036
	(0.126)	(0.172)	(0.0488)	(0.0644)	(0.0424)	(0.0612)
Panel D: *No Cartel-Heavy Districts*						
ddd	-0.0319	-0.244	-0.0942*	-0.112	0.120***	0.00706
	(0.143)	(0.175)	(0.0533)	(0.0701)	(0.0465)	(0.0619)
N	1687	1687	6493	6493	4718	4718

Notes: This table considers four different robustness checks of the the triple-differences analysis, shown across four panels. Panel A contains only crack cases, while Panels B, C, and D contain all drug types. The ddd coefficient is the triple differences effect, representing the interaction of minority, conspiracy circuit, and post. All specifications include defendant controls, population controls, and district and monthly fixed effects. All standard errors are clustered at the district level.

Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Wider drug weight windows

	(1) bunch	(2) MM app.	(3) bunch	(4) MM app.
Panel A: Fair Sentencing Act				
ddd	-0.0232	-0.459***	-0.00814	-0.245***
	(0.0491)	(0.116)	(0.0181)	(0.0728)
N	4309	4309	9306	9306
Panel B: Minus-2 Amendment				
ddd	-0.0392	-0.0711**	-0.0189*	-0.0376*
	(0.0235)	(0.0363)	(0.0111)	(0.0200)
N	13489	13489	32435	32435
Panel C: First Step Act				
ddd	0.0784***	-0.0132	0.0378***	-0.0253
	(0.0241)	(0.0399)	(0.0131)	(0.0248)
N	9109	9109	20609	20609
Drug-weight in %	20-200	20-200	0-300	0-300

Notes: This table replicates the main MM application analysis allowing for a larger window of weights. Each panel reports separate analyses for each policy. Panel A contains only crack cases, while Panels B and C contain all drug types. Columns 1 and 2 give all cases between 20 % and 200 % of the 10-year MM threshold weight. Columns 3 and 4 contain all cases with weights at or under 300% of the threshold weight. All specifications contain defendant characteristics, population controls, and district and monthly fixed effects. Standard errors are clustered at the district level.

Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$