

# How much do Mandatory Minimums Matter?

Spencer Cooper\*

November 15, 2025

## Abstract

Mandatory minimum sentencing is frequently identified as a potential driver of long sentences among drug offenders. I estimate the causal effect of mandatory minimum (MM) eligibility on federal drug sentencing using a regression discontinuity design with extrapolation to disentangle statutory impacts from prosecutorial selection. I find that MM eligibility increases sentence length uniformly across case types by about 10 months (14%). This includes defendants with low criminal history, indicating limited protection for low-level offenders. To assess which types of cases are affected by selection, I compare extrapolated counterfactual sentence lengths against observed sentences. I find evidence that charging manipulation is localized among minority defendants. These results indicate that racial disparities in MM sentencing are driven by prosecutor charging decisions rather than by features of the MM statute.

---

\***Affiliation:** University of Connecticut, Department of Economics.

**Thanks:** I thank Evan Taylor, Katherine Barnes, Tiemen Wouterson, Daniel Herbst, Hidehiko Ichimura, Jason Kreag, Stephen Ross, David Simon, Kevin Schnepel, Eric Helland, and several anonymous assistant US Attorneys and Assistant US Attorneys for helpful comments and insight.

The federal prison system is overcrowded and extremely costly. The Federal Bureau of Prisons estimated federal prisons operated at 10% overcapacity in 2024, with the average cost of incarceration estimated to be \$43,836 per inmate, per year (2023 United States Government Publishing Office). This large number of inmates is primarily comprised of drug offenders, who make up around 45% of all federal prisoners. Drug offenders bring especially high costs due to their long incarceration spells; in 2019, federal drug offenders had sentence lengths 252% higher than non-drug defendants. This was not always the case; both the volume of drug convictions and attached sentence lengths have dramatically increased since the 1980s and 90s (The Pew Charitable Trusts 2015). The cause of this increase is often attributed to legislation that increased punishment schedules for drug offenders. The most commonly cited and controversial aspect of such legislation is mandatory minimum (MM) sentencing. Yet, despite its prominence and controversy in drug reform debate, very little research has been done to estimate the causal impact MM eligibility has on sentence lengths.

Mandatory minimum sentencing restricts a judge's ability to give a sentence length below a specified amount. In practice, over 97% of all trafficking cases are resolved by plea bargain but MM sentencing is still expected to increase sentence lengths by increasing prosecutor bargaining power. MM sentencing applies to several crime types in the federal system, including weapon crimes, sex crimes, certain economic crimes, and most commonly, drug crimes. Federal drug trafficking offenses may be eligible for a 5-year or 10-year MM sentence depending on the drug type and the charged quantity of drugs. Many cases are affected by this practice; in 2019 when looking at the five most common drug types, over 86% of cases were eligible for a MM sentence, with over 69% of cases meeting the threshold

for a 10-year MM charge.

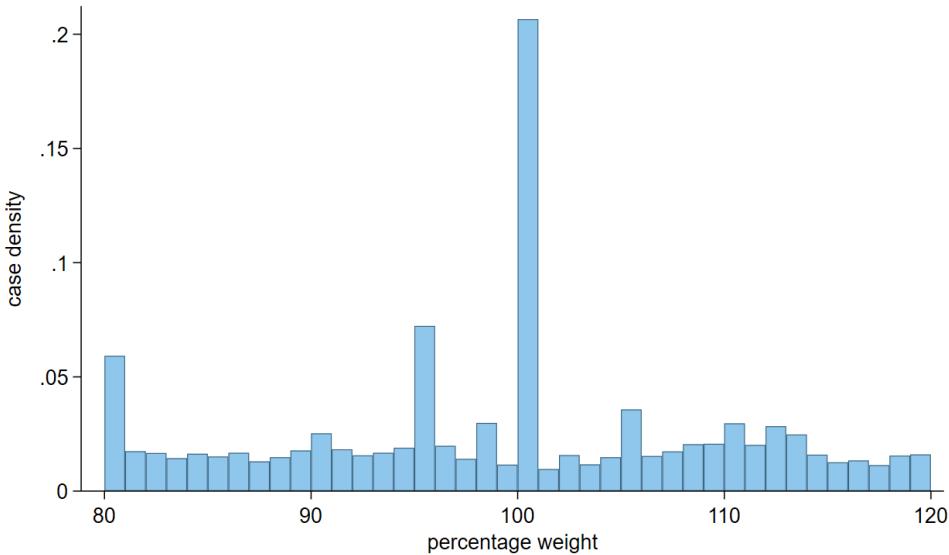
Mandatory minimum laws also raise equity concerns. One of the most common complaints over MM legislation is that it is believed to disproportionately impact racial minority defendants. These concerns are supported in empirical research; Rehavi and Starr (2014) attributes significant disparities in sentence lengths to MM charges and Tuttle (2025b) finds Black defendants are disproportionately targeted for MM sentencing in crack cases. I add to these findings by considering whether eligibility disproportionately increases sentence lengths for minorities.

The primary challenge in estimating the causal impact of MM laws on sentence length comes from selection issues. For most cases, MM eligibility is determined by the drug quantity listed at sentencing, measured by weight. If the weight at sentencing is greater than some threshold value, an individual is eligible to be charged with a 5 or 10-year mandatory minimum. However, this quantity does not need to equal the weight recorded at arrest. Essentially, this means that cases near the MM threshold may have their charged quantity manipulated in such a way that their eligibility is determined in part by legal actors. Tuttle (2025b) notes the presence of significant bunching in the number of crack cases right above the 10-year MM threshold weight, evidence of this manipulation. This means that certain cases are systematically moved from 5-year MM eligibility to 10-year MM eligibility, thereby distorting the relationship between charged weight and sentence length.

I likewise find clear evidence of drug weight manipulation across the five primary drug types. Figure 1 presents a histogram of cases by drug weight for observations near the statutory threshold. The distribution exhibits a pronounced spike precisely at the threshold

weight, providing clear visual evidence of case selection. Such clustering is unlikely to occur naturally, as rational drug offenders would be expected to carry amounts just below the threshold to maximize quantity while avoiding exposure to a mandatory minimum penalty. Notably, this pattern holds consistently across racial groups and drug types, suggesting that manipulation is not limited to a particular subset of cases but instead reflects a broader, systemic feature of federal drug prosecutions.

**Figure 1:** Bunching in drug weights



*Notes:* This figure gives case density across drug weight, normalized to the percent of the 10-year MM threshold weight. The bunching point at 100 indicates a disproportionate number of cases with a weight that is exactly equal to the 10-year MM eligibility cutoff weight, such that these cases are just eligible for the 10-year mandatory minimum.

This paper has two main objectives: first, to estimate the causal effect of 10-year mandatory minimum (MM) eligibility on federal drug trafficking sentence lengths; and second, to identify which types of defendants are more likely to be manipulated into higher MM eligibility. I explore how these effects vary by race, criminal history, and drug type. To

identify causal impacts, I estimate the conditional expected sentence lengths as a function of drug weight using data from unmanipulated regions on either side of the 10-year eligibility threshold. I then extrapolate these estimates into the manipulated regions to construct counterfactual sentence length distributions under the scenario where no manipulation occurs. The discontinuity between these counterfactual trends at the threshold captures the effect of MM eligibility absent manipulation.

I next consider how these counterfactual sentence lengths compare to actual sentence lengths within the manipulation region. Doing so gives insight into the types of defendants that prosecutors are selecting. Specifically, I look near the MM eligibility threshold for large deviations in observed sentences compared to the fitted sentence lengths across different types of defendant. Significant discrepancies between the two provide evidence of prosecutor manipulation.

I find that eligibility for a 10-year mandatory minimum causally increases sentence lengths by about 10 months (or about 14% above the mean). This effect is nearly identical across defendant race, suggesting that observed racial disparities do not stem from the MM statute itself. I also find that MM eligibility produces causal increases for different criminal history levels and across each major drug type. The estimate is smaller, but still present, for first-time offenders. These patterns indicate that protective policies—most notably the safety valve provision—may not fully insulate low-level offenders from the effects of mandatory minimums. Because most of these cases are resolved through a plea bargain, the mechanism for this effect is likely the enhanced bargaining leverage afforded to prosecutors. These causal effects are significant and robust to a battery of alternative specifications and various checks.

Beyond estimating the average effect, I also find significant evidence that certain cases are manipulated from 5-year MM eligibility into 10-year eligibility. The large gap between the predicted sentences and observed sentences indicates that the bunching of cases at the threshold weight is not merely due to round-number bias, but is at least partly driven by prosecutor discretion. Cases most likely to be manipulated to the threshold weight are the “most severe” ones: those going to trial, cases involving a firearm, and cases with high criminal history defendants. Strikingly, these selection patterns appear only for minority defendants. Even when conditioning on the most severe cases, I find no evidence of manipulation for white defendants, whereas significant distortions emerge in the sentence length distributions for minority defendants. Taken together, the results suggest that MM sentencing impacts race gaps through prosecutor charging decisions rather than through the impact of the statute itself.

This paper contributes to literature on the welfare impacts of sentencing structure, legal actor discretion, and how each of these may drive racial disparities. Many papers suggest legal actors use discretion to disproportionately target or punish racial minorities with worse court outcomes (Arnold et al. 2018; Rehavi and Starr 2014; Sloan 2022; Tuttle 2025b; Yang 2016), though recent literature has suggested prosecutor discretion is not a driver of race gaps along certain dimensions and prosecutors may actually limit racial disparities, specifically in connection with enforcement decisions (Shaffer and Harrington 2017; Shaffer 2023; Yuan and Cooper 2022). Findings from this paper support previous results; that race disparities are largely driven by discretion and are likely tied to prosecutor choices. On a smaller scale, this paper also adds to the drug crime, MM sentencing literature. Findings among these studies vary, with some papers downplaying the significance of mandatory

minimums (Bjerk 2005; Bjerk 2017a; Bjerk 2017b; Fischman and Schanzenbach 2012) and others finding significant impacts on sentencing generally and in contributing to racial disparities (Didwania 2020; Didwania 2025; Rehavi and Starr 2014; Tuttle 2025b). I find a nuanced impact of MM sentencing, where eligibility has a significant increase to sentence lengths, but with magnitudes smaller than previous papers. Likewise, I find that race is not causally impacted by eligibility directly, but that only minority defendants show evidence of being moved from 5-year to 10-year MM eligibility.

Bjerk (2017b) is the paper most closely related to this work. It provides a detailed look at which cases are eligible and which receive MM sentencing, with a specific focus on crack cases. Most relevant for this study, Bjerk (2017b) similarly considers MM eligibility impacts on sentence lengths but does so only for cases very close to the threshold weight and only for fiscal years 2011 and 2012. Thus, Bjerk's analysis is a traditional regression discontinuity design that includes selection bias, which is explicitly acknowledged in the paper. Indeed, I find that because the regression in Bjerk (2017b) fits on cases most likely to be manipulated, the magnitude of the eligibility effects are significantly inflated; over twice as large as the causal effects I estimate. Beyond accounting for selection bias and increasing the time frame of the study, I also contribute to Bjerk (2017b) by highlighting which types of cases are most likely to be manipulated. This provides a unique look at how MM eligibility impacts race disparities, showing consistent causal impacts across race but disproportionate selection for minorities to be moved to the threshold weight.

## 1 Background

### 1.1 Mandatory Minimum Sentencing of Federal Drug Cases

The main criteria for MM eligibility is drug quantity.<sup>1</sup> In order for the mandatory minimum to apply, the weight at sentencing for one drug type must meet or exceed the set threshold weight. Weights cannot be added across drug types in regard to MM eligibility, meaning there must be a large quantity of at least one drug type. There are two separate thresholds, one for a 5-year and one for a 10-year mandatory minimum. In this paper, all analysis considers only the higher threshold, which has stronger bunching, more severe punishment increases for eligible cases, and for which 70% of all federal trafficking cases are eligible.

Cases with drug quantities at or above the MM threshold weight may not necessarily be charged with a mandatory minimum. But being charged at an eligible weight opens the possibility for the prosecutor to impose a MM charge, meaning hitting the threshold weight significantly increases prosecutor bargaining power. Specifically, for a MM eligible case a prosecutor may use the threat of a MM sentence should the case go to trial, allowing them to secure a higher sentence length in plea negotiations. Other cases may have initial charges that apply the MM filing, but through plea negotiations, do not carry a mandatory minimum in the final charges. In my data, I only observe final charges. For this reason, I do not focus on the MM charges themselves but only consider drug quantity and the

---

<sup>1</sup>A small subset of cases are eligible for MM penalty without meeting quantity thresholds. This can occur in two ways: if the crime involves death or serious injury, or if the defendant has committed a serious prior drug offense. Including, excluding, or controlling for this subset does not substantially change results or significance of findings.

sentence length. This ensures I do not leave out cases that do not show MM charges in the final charge data, but were still substantially impacted by MM eligibility during the plea bargaining process.

Furthermore, MM sentences are often non-binding, with many convictions receiving sentences above or below the minimum specified sentence length. Sentences below the mandatory minimum can occur if a defendant provides “substantial assistance” in the prosecution or investigation of another offender,<sup>2</sup> if the offender is eligible for what is called the safety valve provision, or if a lower sentence is negotiated through plea bargain. Defense can apply for the safety valve provision if the defendant in question meets 5 specific criteria: they have a sufficiently low criminal history record, they did not have a leadership role, they did not use violence or possess a firearm during the crime, no persons were harmed, and the defendant agrees to truthful disclosure. If these conditions are met, the provision allows the courts to assign a sentence lower than the prescribed MM sentence length. I directly test for MM effects for cases filed under the safety valve provision.

## 1.2 Drug Weight Manipulation

Discrepancies between seized, charged, and sentencing drug weights may occur for several reasons, including round-number bias, converting seized cash to drug weight, or leniency. A further source of manipulation—and the one most relevant for this study—arises from changes in the evidentiary record. This may occur through additional testimony or by connecting a defendant to other traffickers or cases (Lynch 2016). The impetus for this additional evidence comes from prosecutors, who decide whether to pursue further investigative links when building a case. If prosecutors can connect a defendant to other offenders or

---

<sup>2</sup>See USSG § 5K1.1

larger organizations, they may increase the available evidence of drugs to charge a person with. Importantly, manipulation through this evidentiary channel almost always pushes cases upward—moving them from lower weights up to the threshold weight.

One indication of this evidentiary channel appears in conspiracy charges. 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. Thus, if prosecutors seek to connect defendants to other offenders or an organization, they are more likely to charge them with conspiracy. Tuttle (2025b) finds evidence of this in crack cases, showing a high concentration of conspiracy charges for cases at the threshold weight. This appears true among other drug types as well; I find that for my full sample of the five main drug types, 54% of cases that are not at the bunching point carry a conspiracy charge. However, at the MM threshold weight, 83% carry a conspiracy charge. Conspiracy rates are also comparable across racial groups, with White defendants having slightly higher rates than non-White defendants.

Drug weight manipulation is an important feature to account for in this analysis for two reasons. The first is that it poses an empirical challenge in uncovering the causal effect of MM statutes separate from selection bias. The second is that manipulation reveals which types of defendants are disproportionately selected into 10-year MM eligibility. In this paper, I am specifically interested in which types of cases prosecutors are moving upward from 5-year MM eligibility into 10-year MM eligibility. However, by just looking at the distribution of cases, it is nearly impossible to determine which cases are manipulated rather than merely bunched for unrelated reasons. The sentence-length analysis I provide offers evidence that a substantial number of cases are indeed manipulated from below and

identifies the types of cases most likely to be shifted from 5-year to 10-year MM eligibility.

## 2 Data & Empirical Strategy

### 2.1 Data

The primary drug case data is provided by the United States Sentencing Commission (USSC) and includes all federal drug trafficking cases from 2010 to 2021. Data is at the case-individual level. I restrict the data to the five most prevalent substances subject to MM sentencing: powder cocaine, crack, heroin, methamphetamine, and marijuana.<sup>3</sup> I also restrict the data to White, Black, and Hispanic defendants. Finally, the data is further restricted to include only cases with primary drug weights at 20% to 180% of the threshold weight. This gives enough data to fit distributions on but excludes the 5-year threshold and extremely high weight cases that are less similar to cases near the 10-year threshold. This gives a total of 44,626 observations.

The USSC data provides a rich set of defendant and litigation details. Information on the defendant's sex, race, age, education, and citizenship is included. The data also contains specific statutes that are charged, the drug weight at sentencing, any factors that increased or decreased the sentence length above or below the prescribed guideline amount, whether the case was plead or not, and sentence length the defendant received.

Each of the five drug types have a sizable number of cases, with heroin cases making up the fewest percent of cases at 14.3% and cocaine making up the most with 26.1%. Drug type is highly correlated with race, suggesting primary specifications for race heterogeneity should control for the primary charging drug. The table also shows that Black and White

---

<sup>3</sup>The other three substances subject to MM law are PCP, LSD, and fentanyl. These had too few observations for any meaningful analysis.

defendants have very similar criminal history points, while Hispanic defendants have much lower criminal history on average. I also consider conspiracy charges as these are a primary mechanism by which drug weight manipulation occurs (Lynch 2016, Tuttle 2025b, Cooper 2023). Across all three racial groups, the proportion of cases with a conspiracy charge is nearly identical. This helps alleviate concerns of unequal opportunity for manipulation across races. These and other summary statistics are shown in Appendix Table A.1.

## 2.2 Empirical Strategy

To accurately estimate the causal effects of MM sentencing, I need to know what the discontinuity at the MM threshold looks like without any manipulation. Thus, my strategy is to create fitted correlations between sentence lengths and charged drug weight over the range of unmanipulated drug weights. These fits are then extrapolated into the regions near the eligibility cutoff where manipulation is present. The extrapolated fits serve as counterfactuals under the scenario where manipulation does not occur. This technique is a regression discontinuity design using extrapolation, often referred to as a “donut RD”. I fit two separate distributions on either side of the cutoff. The fitted distributions are determined only by cases that are assumed to be unmanipulated. I then look at the extrapolated points right at the threshold weight. The distance between these two points gives the effect of MM eligibility in absence of manipulation activity.

The model can be expressed with the following notation: let  $r_i = \text{percent\_weight}_i - 100$  denote the running variable, drug weight as a percentage of the threshold weight, centered at the cutoff, 100. Let  $D_i = \mathbf{1}\{r_i \geq 0\}$  be an indicator for being to the right of the cutoff. Let  $h_L$  and  $h_R$  represent the bandwidth ends for which extrapolation start from for the left

and right sides of the cutoff, respectively. Similarly,  $\delta_R$  and  $\delta_L$  represent the points at which extrapolation begins from each side of the cutoff. Using only observations that satisfy

$$h_L \leq r_i \leq \delta_L \quad \text{or} \quad \delta_R \leq r_i \leq h_R,$$

I estimate the following specification:

$$\text{Sentence\_Length}_i = \alpha + \beta_1 D_i + \beta_2 r_i(1 - D_i) + \beta_3 r_i D_i + u_i. \quad (1)$$

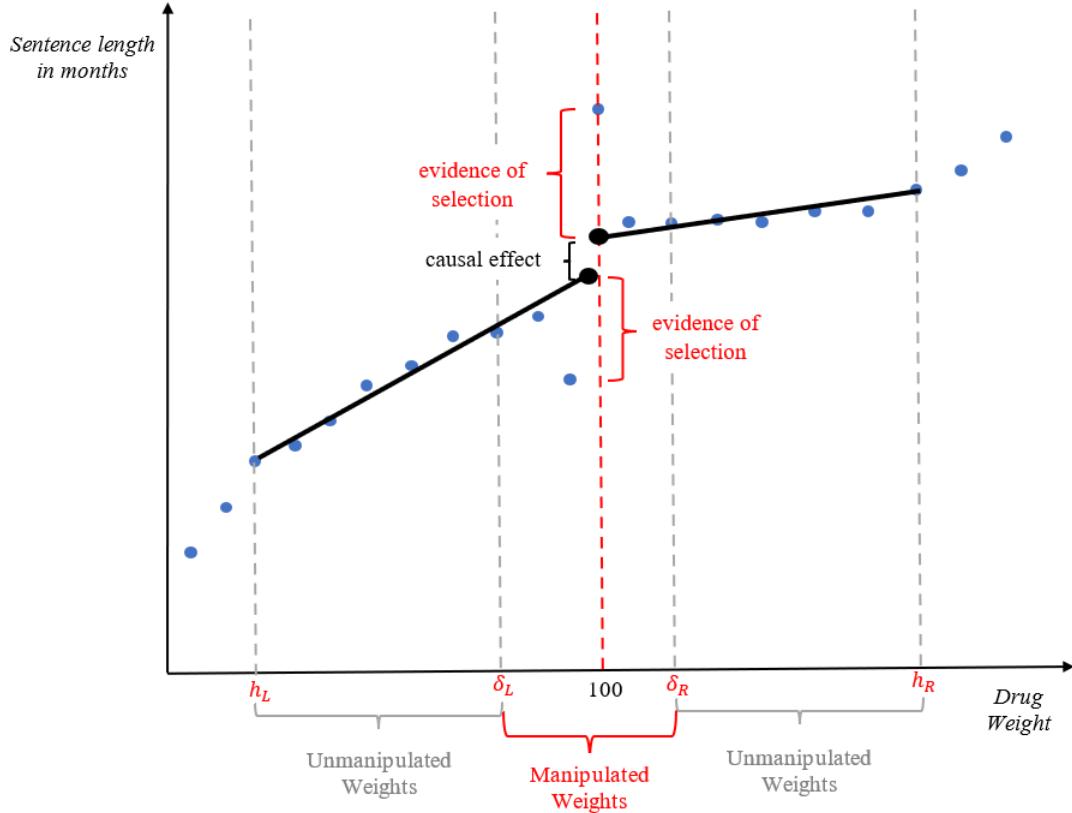
where  $\beta_1$  captures the estimated discontinuity at the cutoff via the extrapolated linear fits from each side to the cutoff at  $r_i = 0$ .

I also consider how these extrapolated regressions compare against the actual sentence lengths. I compare the point extrapolated from the right-hand side at the threshold weight against the actual sentence length at the threshold weight. Essentially, I am comparing the counterfactual sentence length at the threshold weight versus the observed sentence length at this point. I also look for a discontinuous drop in sentence length for cases just before the cutoff. If prosecutors manipulate cases that are already likely to have higher sentences, the “most severe cases,” then the average sentence where they are manipulated from will be lower than the rest of the trend. Note that this is analogous to a missing mass argument.

Figure 2 provides an illustrative example of this strategy. Note that this figure does not contain actual data but is simply meant to conceptualize the main models used in my analysis. The blue points represent the actual sentence length distribution for the group being considered. Dashed vertical lines show the manipulated and unmanipulated regions. The black lines give the fitted distributions which are extrapolated into manipulation regions. The large black points give the discontinuity absent manipulation, the causal effect of MM eligibility. Large deviations from the black lines in observed points give evidence of

selection by prosecutors.

**Figure 2:** Empirical strategy illustration



*Notes:* This figure depicts the empirical strategy used to estimate the MM causal impacts and the way I consider selection impacts. The black lines represent the fit over the unmanipulated weights and extrapolated into the manipulated weights. The discontinuity between these fits at the threshold is the causal effect. The gap between the predicted sentence from the right-hand side of the distribution and the actual sentence length is the primary consideration in the selection patterns analysis.

To implement this empirical strategy, I must first determine suitable values for  $\delta_L$  by finding which regions are manipulated and which are not. There are two primary ways in which manipulation regions have been detected and determined in past literature. The first is a formal test developed in Frandsen (2017) and practiced in Goncalves and Mello (2021) used to detect changes in the distribution attributed to manipulation. This method

is unlikely to work well in this setting because small scale manipulation likely occurs as prosecutors or law enforcement round to whole numbers. The second method is simply using visual inspection, as done in Saez (2010), Chetty et al. (2011), and Kleven and Waseem (2013). This approach works when there is a clear and obvious missing mass in the distribution that is supplying the observations at the bunching point. Where the missing mass begins is assumed to be the beginning of the manipulation region. To tease out the missing mass area, I fit a fifth order polynomial with fixed effects for each 10-percentage point round figure over the main analysis weights: 20% to 180% of the threshold weight. Appendix Figure A.1 shows this fit distribution in comparison to the actual case density. Missing mass appears following the 70% value and continues up until cases just before the MM threshold at 100%.

One common way to check whether the manipulation region is correctly specified is to compare the excess and missing mass amounts, which should be equal. The excess mass at the bunching point is just under 2,385 cases. The missing mass from 70% to 99% is about 1,129. The smaller missing mass may be generated by two sources. First, when looking at each drug individually, meth has a manipulation range that is clearly wider than the other drug types (closer to 60%). Second, the missing mass estimation does not include cases at 70%, 80%, or 90%. These round numbers are omitted as rounding bunching occurs at these points as well. Significant case reduction may occur from cases at the 80 percentage point cases.

Based on the comparison between the fit and actual case distribution, I consider cases between 70% and 105% to be within manipulation ranges for my main analysis. This means the left-side regression is fit using cases between 20% and 70% of the threshold

weight, while the right-side regression is fit using cases between 105% and 180%. Each of these fit predictions are then extrapolated into the manipulation region to create counterfactual distributions, or the trend of sentence length absent manipulation. In my robustness checks I consider the results under many specifications with different cutoffs for the left-side manipulation region. There is a tradeoff in setting the manipulation region cutoff; cutoffs further from the threshold are less likely to be biased since they are less likely to accidentally include cases that are manipulated, while windows closer to the threshold have more data to fit on and are thus likely to be more precise. For this reason, it is important to show that results are similar across many different set cutoffs. This is analogous to testing a regression discontinuity with different bandwidth sizes.

Following Gelman and Imbens (2019), my analysis assumes a functional form that is linear, though I include some quadratic fit predictions in my robustness checks. I use standard errors of the prediction to create confidence intervals for each fitted value to assess inference. All tables with extrapolated RD present confidence intervals on either side of the cut-off point.

The key identifying assumption of this empirical design is that without manipulation, the distribution of sentence lengths for cases in manipulation regions would have followed the counterfactual distributions fit using unmanipulated cases. Note that this assumption is standard in a traditional bunching design, only there it is typically involving the variable being bunched, which in this context would be cases. Now I assume that the relationship between sentence length and charging weight can be predicted using the polynomial coefficients among unmanipulated cases.

The main methodological concern is that prosecutors may manipulate cases with lower

drug weights (0–50% of the threshold) up to the bunching point. However, such manipulation is unlikely for four reasons. First, manipulation typically requires additional fact-finding by prosecutors (Lynch, 2016; Tuttle, 2025b), and the cost of gathering such evidence increases with the amount manipulated (Tuttle, 2025a). If the availability of evidence is roughly uniform across seized weights, manipulation should be most feasible just below the threshold and less likely at lower weights. Second, I test six other covariates and outcomes for discontinuities at the threshold and find no evidence of breaks; this analysis is discussed in the robustness section. Third, missing sentence lengths cluster just below the threshold, consistent with strategic omission rather than widespread manipulation across lower weights. Finally, in the robustness section I also show that the RD estimates hold even in districts with minimal bunching, serving as a counterfactual where manipulation is less likely.

### 3 Results

#### 3.1 Causal Effects

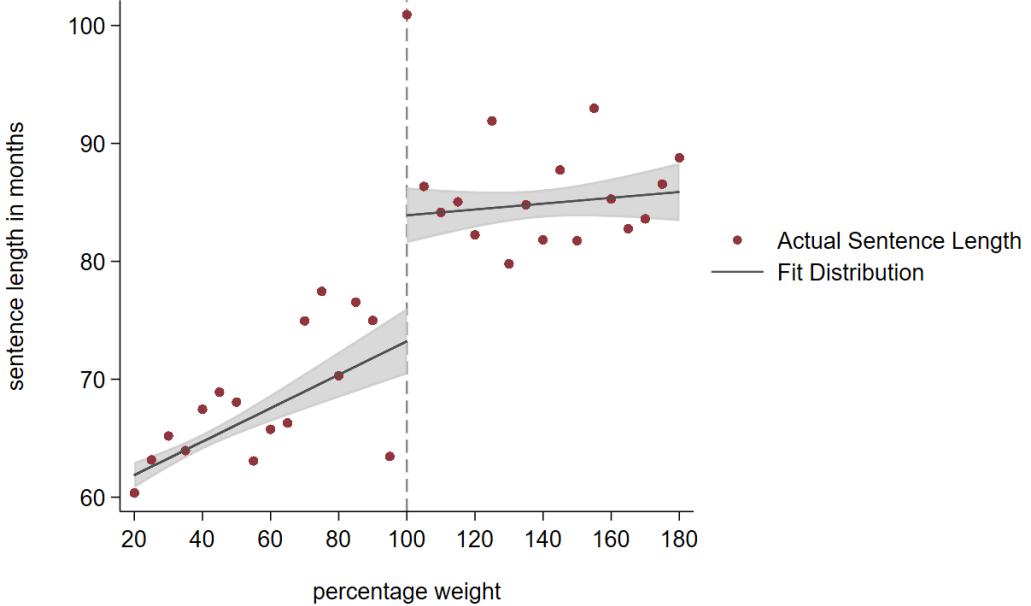
Figure 3 displays the aggregated discontinuity across all cases. MM eligibility is shown to increase sentence lengths by 10.63 months — a 13.9% rise over the predicted mean at the 99% threshold weight. The fitted values closely track actual sentence lengths within the manipulation range, except for cases nearest the threshold, suggesting manipulation is concentrated around that point. Estimates are robust across a range of control specifications, including defendant and case characteristics, district and year fixed effects, and prosecutor decision variables.<sup>4</sup> The estimated 10-month discontinuity remains stable across all models.

---

<sup>4</sup>Baseline controls include criminal history points, drug type, race, age, education, and gender. Prosecutor decision controls include a gun use indicator, the number of drug types charged, and whether the case

Full results are reported in Appendix Table A.2, with robustness checks using alternative fits shown in Appendix Table A.3.

**Figure 3:** Discontinuity analysis for all cases



*Notes:* This figure gives the main discontinuity analysis for all cases with linear fits through points before 70% of the threshold weight and extrapolated past 70% to the threshold from either side. Sentence lengths are binned by 5% except at the threshold weight, which only reports average sentences for cases charged with exactly the threshold weight.

### 3.1.1 Traditional RD

I also estimate the discontinuity under a traditional design—that is, without excluding cases that may be biased. The results, reported in Appendix Table A.4, indicate that failing to omit cases near the threshold leads to substantially inflated estimates. Across five different bandwidth specifications, the estimated discontinuities range from 21.4 to 40.7 months, all of which are considerably larger than those obtained from specifications that went to trial.

extrapolate across manipulated cases. This analysis essentially replicates the approach of Bjerk (2017b), and the magnitudes are correspondingly close to the 24.6-month estimate reported in that study. These findings underscore the importance of excluding cases near the threshold to obtain unbiased estimates of the effects of MM eligibility.

### **3.1.2 Race, Criminal History, and Drug Type Heterogeneity**

Next, I examine how the results vary by defendant race, criminal history, and primary drug type. Discontinuity estimates for these subgroups are presented in Table 1. The estimates are nearly identical across racial groups, suggesting that mandatory minimum (MM) eligibility does not affect sentence lengths differently by race—at least on the intensive margin. This finding does not rule out the possibility that prosecutors apply MM eligibility criteria differently across racial groups, an issue I address later in the discussion of selection patterns. Rather, it indicates that, conditional on eligibility, a defendant’s race does not causally influence the sentence length they receive.

To examine MM eligibility effects on low-level offenders, I focus on two subgroups: defendants with no prior criminal history, and those with some prior contact but only 0 or 1 criminal history point, corresponding to offenses with sentences under 60 days. I then compare these groups to the rest of the sample. Because there are policies aimed at protecting first-time and low-level offenders from MM sentencing, the expected effect of MM eligibility for these groups should be small. Indeed, Bjerk (2017b) reports almost no effect of MM eligibility for low-level defendants. In contrast, I find significant discontinuities for both no-history and low-history offenders. These effects are smaller than those observed among defendants with more extensive criminal histories, but nonetheless suggest that existing

protections do not fully insulate low-level offenders from MM eligibility effects. Graphs for low-level offender analysis are presented in Appendix Figure A.3.

When disaggregated by drug type, the effects show greater heterogeneity. The RD coefficients range from 16.07 months for methamphetamine cases to 11.06 months for crack cases. It is somewhat surprising that the estimated impact is the smallest and statistically insignificant for crack cases, given the historical controversy surrounding crack-related mandatory minimum sentencing. However, while crack cases may exhibit smaller discontinuities, they still have the highest overall average sentence lengths. For cases falling between 20% and 180% of the MM threshold, the average sentence length for crack offenses is approximately 100 months; approximately 34 months longer than the average for non-crack cases. Discontinuities by drug type are shown in Appendix Figure A.4.

The results can also be considered at each drug-race intersection. However, these samples become smaller and noisier, leading to less precise estimates. Overall, the pattern remains consistent: estimated effects are broadly similar across racial groups but tend to be largest for the group that comprises the majority of cases within a given drug type. In some cases, one or two racial groups account for nearly all observations for a particular drug, for example black defendants in crack cases. In these settings, discontinuities appear only within the predominant group.

### 3.2 Selection Patterns

I now focus on patterns of selection by comparing cases within the manipulation region to their conditional expected sentence lengths. Although this analysis is descriptive rather than strictly causal, it sheds light on how prosecutors decide whom to manipulate. In particular,

**Table 1:** MM effects heterogeneity

<i>Panel A: Race</i>				
	<u>black</u>	<u>hispanic</u>	<u>white</u>	
RD estimate	11.87	12.03	11.71	
left side 95% CI	[-0.04, 9.94]	[-1.91, 3.25]	[-6.74, 2.36]	
right side 95% CI	[12.28, 21.96]	[10.32, 15.38]	[5.98, 13.38]	
	{11205}	{15296}	{7190}	

<i>Panel B: Criminal History</i>				
	<u>none</u>	<u>low history</u>	<u>med-high</u>	
RD estimate	7.66	9.98	13.45	
left side 95% CI	[0.12, 5.78]	[-1.77, 3.84]	[-3.68, 8.43]	
right side 95% CI	[7.76, 13.74]	[8.29, 14.12]	[10.90, 21.64]	
	{5069}	{9801}	{18821}	

<i>Panel C: Drug Type</i>					
	<u>cocaine</u>	<u>crack</u>	<u>heroin</u>	<u>marijuana</u>	<u>meth</u>
RD estimate	14.56	11.06	15.17	13.42	16.07
left side 95% CI	[-5.49, 2.55]	[3.54, 17.92]	[-7.066, 4.721]	[6.13, 13.22]	[-8.15, 1.65]
right side 95% CI	[9.30, 17.26]	[14.33, 29.76]	[10.58, 20.61]	[17.32, 29.30]	[8.15, 17.68]
	{8484}	{6358}	{4303}	{6287}	{8259}

*Notes:* All specifications are discontinuities based on local linear fits. 95% confidence intervals are presented for each fit regression on either side of the cutoff. These are calculated using the standard error of the predicted expected value. Number of observations used for the fit are in curly braces. Regressions control for race, criminal history, drug type, an illegal alien binary, and a college binary, though omitting any control when it is the dependent variable. For each specification, left-hand regressions are fit on cases with weights between 20% and 70% of the threshold weight and right-hand regressions use weights between 105% and 180%.

by examining deviations in actual sentence lengths and other observable outcomes from their predicted counterfactuals, I can identify which types of defendants prosecutors tend to push into the higher MM eligibility region.

As shown in Figure 3, there are large distortions in sentence lengths near the threshold weight. This pattern highlights three key points. First, bunching in the drug weight distribution is not merely the result of round-number bias; instead, it reflects prosecutors' active selection of certain defendants for higher eligibility. Second, the evidence suggests

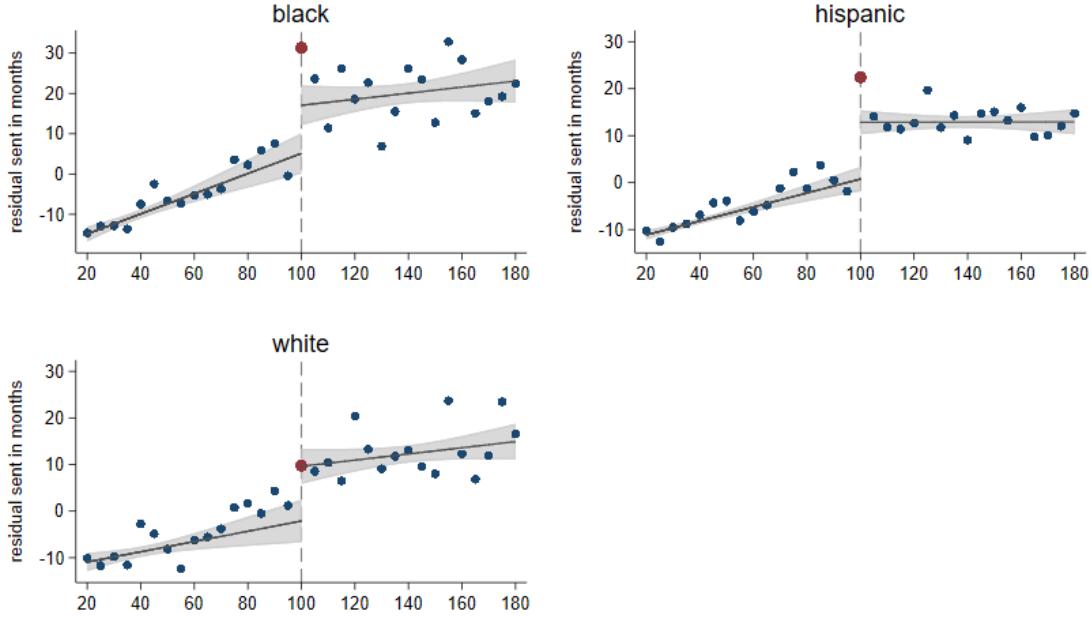
that manipulation primarily affects cases that would have received harsher sentences even without manipulation. Finally, it indicates that manipulation is concentrated among cases very close to the threshold weight.

I find that evidence of selection is only present for the racial minority defendants. Figure 4 shows disproportionately high sentence lengths for cases at the threshold weight for both Black and Hispanic defendants, but no distortion for White ones. This finding is in line with Tuttle (2025b), showing that prosecutors use manipulation in a way that disproportionately impacts minorities. However, Tuttle only considers crack cases. In contrast, I find no evidence of selection for cases with crack as the predominant drug type and only distortions for cocaine and heroin cases.

To further understand which types of defendants are manipulated to the threshold, I now consider selection patterns across other characteristics or outcomes. I focus on six factors - defendant criminal history, whether a gun was involved in the incident, the number of counts in the case, whether the case went to trial, defendant sex, and whether the safety valve provision was applied. Based on the above results, I focus this correlative exercise on minority cocaine and heroin cases, which exhibit the strongest evidence of selection. Panel A of Figure 5a provides graphs of these six variables plotted across drug weight for his subsample of likely manipulated cases. I find evidence of significant selection across these variables with large distortions at the threshold weight. Specifically, cases at the threshold weight have disproportionately high criminal history points, high number of counts, are more likely to involve a gun, are more likely to go to trial, are less likely to be a female defendant, and are less likely to receive the safety valve provision.

Graphs in Panel B of Figure 5a present average sentence lengths by drug weight across

**Figure 4:** Race comparison - discontinuity using residualized sentence lengths



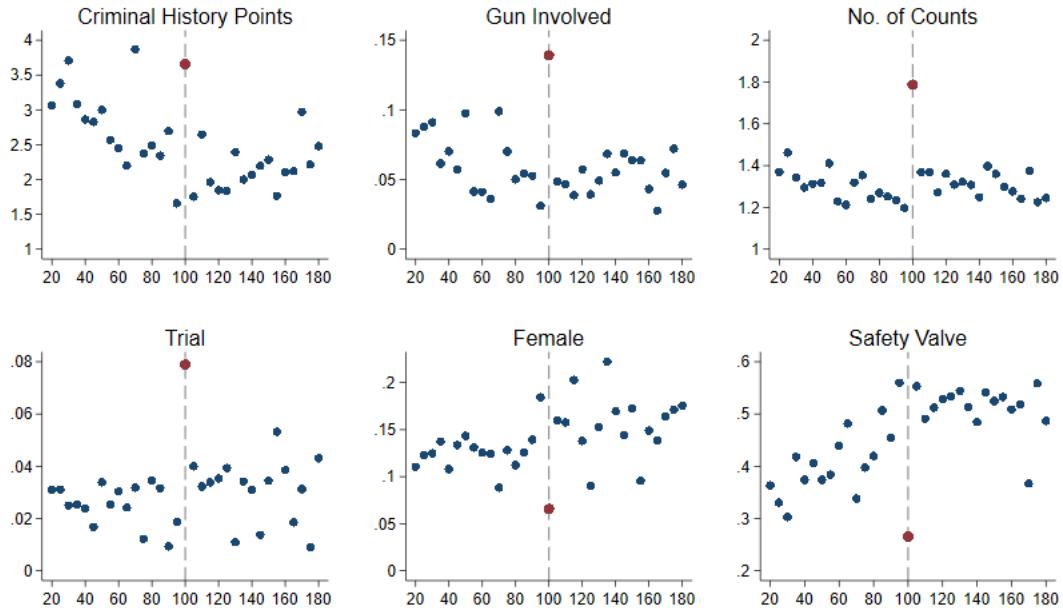
*Notes:* This figure gives the residualized sentence discontinuity analysis by racial group with linear fits through points before 70% of the threshold weight and extrapolated past 70% to the threshold from either side. Residual sentence lengths are binned by 5% except at the threshold weight, which only reports average sentences for cases charged with exactly the threshold weight.

subsets of White defendant cases likely to experience manipulation. The idea is to see if selection patterns arise in sentence length trends for White defendants when conditioning on factors highly correlated with case manipulation. Even within these subsets—those most likely to involve evidentiary or charging manipulation—White defendants continue to show no meaningful distortions in sentence length at the threshold. Put differently, even among the White defendants who should, based on observable characteristics, be the strongest candidates for upward manipulation, their sentence lengths align closely with predicted counterfactuals. This pattern reinforces the conclusion that race is a central determinant

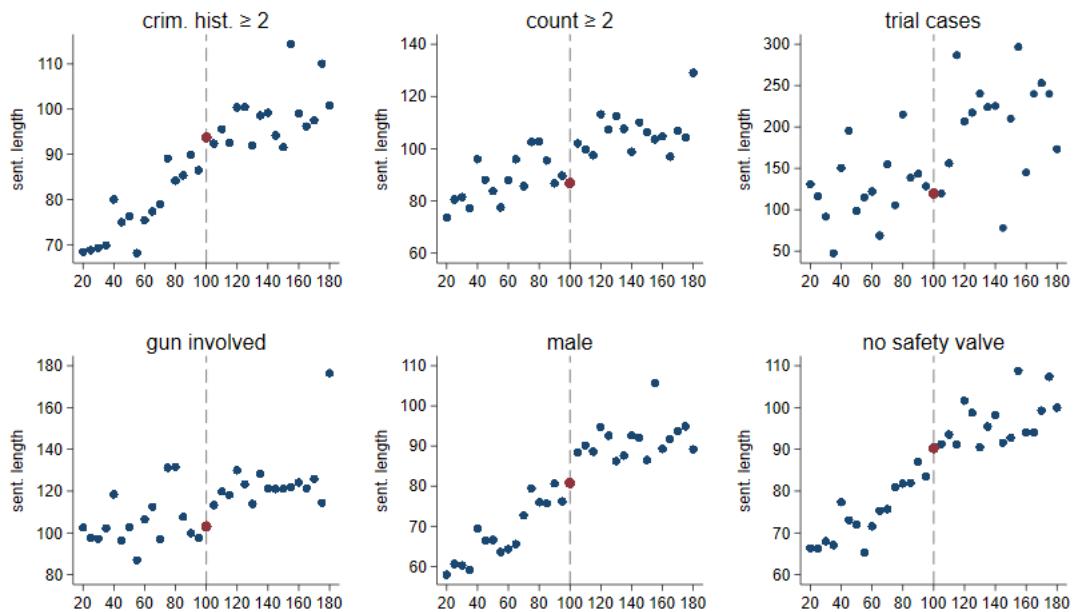
of whether a case is manipulated upward into higher mandatory minimum eligibility.

**Figure 5:** Selection patterns across key case characteristics

(a) Minority cocaine and heroin cases



(b) White cases with high likelihood of manipulation



*Notes:* Panel A gives the average of various case characteristics across drug weight for minority cocaine and heroin cases. That is, for each graph the dependent variable is the proportion of cases with the given characteristic or the average value for the characteristic. Panel B gives the relationship of sentence length across drug weight for White cases that have high likelihood of manipulation based on Panel A results, where each dependent variable is average sentence length.

### 3.3 Robustness

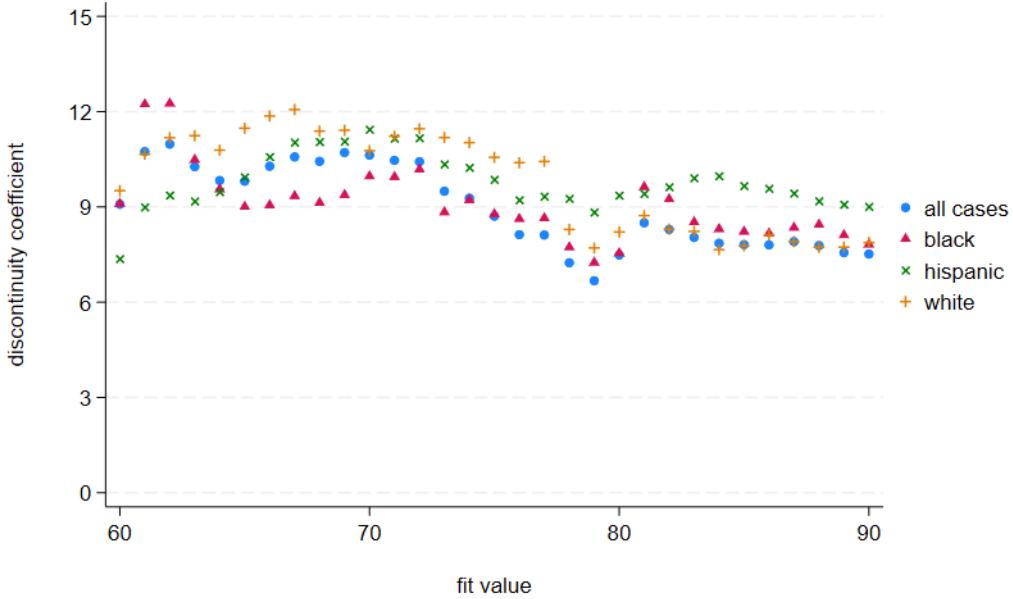
One of the primary concerns for the main results is that using the  $\delta_L = 70\%$  cutoff for fitting over non-manipulated cases either creates bias (if 70% is too low) or doesn't capture all the variation in the data well (if 70% is too high). To check that results and magnitudes are not driven by one specific fit, I repeat the analyses for the overall results and the heterogeneous results using a variety of different fit cutoffs. For each group, I run the main regression discontinuity again for every cutoff between weights 60% and 90% of the threshold for a total of 30 additional specifications (in addition to the 70% analysis).

Results of this test are provided in Figure 6, which gives the discontinuity estimate for all 31 regressions for overall and race-specific effects. Across each specification, I find the main results are supported: MM eligibility causes a significant increase in sentence length, but the magnitude of this increase is considerably smaller than those estimated in Bjerk (2017b). Causal discontinuities stay similar across race groups regardless of which value of  $\delta_L$  is chosen.

Similarly, the main results remain consistent when varying the value of  $\delta_R$ . To account for the possibility that cases may be rounded down on the right-hand side of the threshold, whether due to round-number bias or some form of leniency, I re-estimate the main specifications using five alternative values of  $\delta_R$ , while holding the standard  $\delta_L = 70\%$  fixed. The corresponding results, reported in Table A.6 in the appendix, yield slightly larger magnitudes (ranging from 10.74 to 13.44) but remain broadly comparable to those in the baseline analysis.

To further address concerns about non-random sorting among observations classified as

**Figure 6:** Discontinuity robustness - all cases and by race

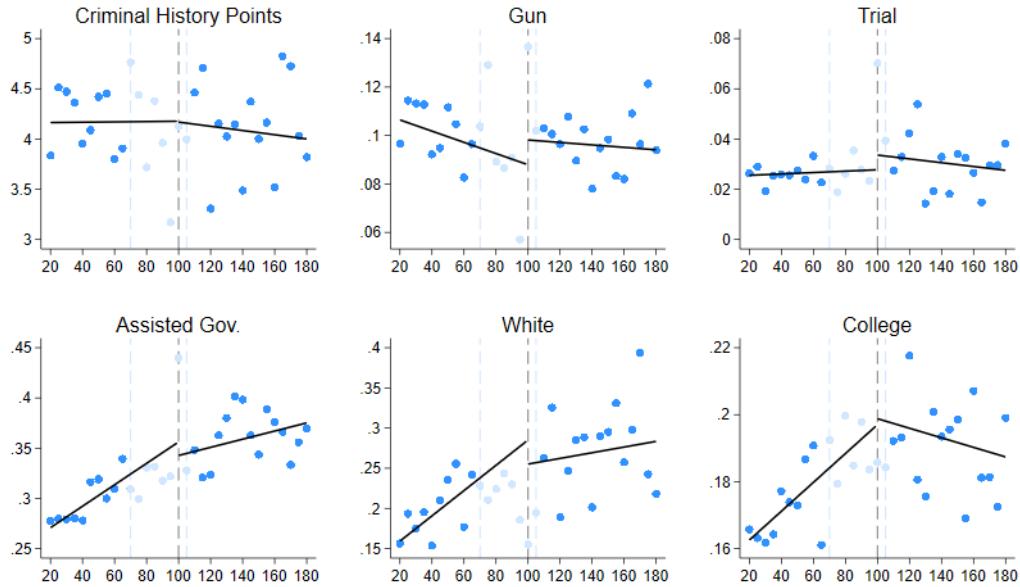


*Notes:* Each point gives the magnitude of the discontinuity under a specific fit value. A fit value designates which area is considered manipulation region and which is considered unmanipulated. For example, at a fit value of  $\delta_L = 65$ , the regression is fit using all cases with weights between 20% and 65% of the threshold weight. The regression is extrapolated from 65% up to the 10-year MM cutoff, where it is compared against the right-hand regression to estimate the 65% discontinuity coefficient.

non-manipulated weights, I conduct smoothness checks on several observable defendant and case characteristics. Consistent with the main analysis, I exclude cases near the threshold weight, as these are most likely subject to manipulation. For each variable, I estimate the extrapolated regression discontinuity using the primary specification—a linear model with bandwidths and extrapolation ranges set to  $h_L = 20\%$ ,  $\delta_L = 70\%$ ,  $\delta_R = 105\%$ , and  $h_R = 180\%$ . The results of these donut RD tests are presented in Figure 7. As shown in the figure, none of the characteristics exhibit meaningful discontinuities, suggesting that sorting among cases further from the threshold weight is minimal.

I also show that the causal effects of MM eligibility are legitimate by comparing them

**Figure 7:** Smoothness across other observables



*Notes:* The following graphs display donut RD analyses across six different case characteristics that may correlate with prosecutor selection. Each regression is fit between 20% and 70% of the threshold weight on the left side of the cutoff and 105% and 180% on the right side of the cutoff.

against districts with low bunching levels. Districts without bunching represent jurisdictions with low or without any manipulation. I compare the main results against a series of district subsets with low bunching amounts. I find the discontinuities in these subsets are consistent with the main results. These are presented in Appendix Table A.7.

I also check that the main results hold using a quadratic fit. For this check, I only use a cutoff value for the manipulation region of 80% of the threshold weight. I use a higher cutoff because at 70%, the quadratic polynomial tends to overfit the data for certain groups. Thus, to assure a reasonable fit I increase the amount of data to fit on and decrease the amount of extrapolation needed. The quadratic fits do not largely change the results, only slightly increasing or decreasing effects.

In the main analysis, I control for the number of other drug types a defendant is charged with outside of the primary drug type. Readers may be concerned that the quantity of other drugs is driving results, especially if other drug type quantities are correlated with certain racial groups. I consider both the causal effects and selection impacts for the full sample and across racial groups again, now controlling for the quantity of other drug types rather than just the number. These results are presented in Appendix Table A.8. I find effect sizes and statistical significance very close to those listed above, though with smaller selection impacts.

Another concern may be that limiting the sample to cases with precise weights biases results if imprecise measures are strongly correlated with sentence length. I now rerun the full sample analysis including cases that are charged with a range of weights rather than one precise count. I use three different measures to do this - the minimum, median, and maximum of the range. These range cases are combined with the precise weight cases for a combined samples of 51,975 cases when using the maximum or median, and 56,189 when using the minimum measure. Regression results for this sample are presented in Appendix Table A.9. Across all three measures, the causal effects and selection impacts have similar magnitudes to main results and remain statistically significant.

## 4 Conclusion

In this paper, I show that mandatory minimum eligibility for federal drug cases has two important impacts on sentencing. First, eligibility for the 10-year MM sentence causally increases sentence length for the average case by 10.63 months, which is a 13.9% increase over the counterfactual mean. Given a defendant is eligible, this effect is effectively the

same across defendant race. MM impacts persist across all drug types, though with smaller effects for crack cases, and still impact defendants with low levels of criminal history. While this effect is significant, the magnitude is substantially smaller than prior literature and certainly much smaller than the 5-year gap between the 5-year and 10-year eligibility levels. Taken together, these results indicate that the statutory enhancement meaningfully affects sentencing outcomes, but the size of the causal effect is far more modest than the nominal increase built into the statute.

Second, MM eligibility creates bunching at the threshold weight, with a disproportionate amount of cases being sentenced right at the 10-year MM eligibility cutoff. While cases may be bunched for several reasons, I provide evidence that at least a portion of these cases are manipulated upward from 5-year to 10-year eligibility. These cases tend to be those with high sentence length correlates, including gun use, high criminal histories, and a high number of counts. Crucially, the cases selected for higher eligibility appear to be almost exclusively minority defendants, suggesting that manipulation does not occur randomly across the defendant population. Taken together, these two findings suggest that race disparities in MM sentencing are driven by prosecutor charging decisions rather than by features of the MM statute. This highlights the central role of prosecutorial discretion in producing unequal sentencing outcomes, even within a rigid mandatory minimum framework.

## References

- Arnold, D., W. Dobbie, and C. S. Yang (2018). Racial bias in bail decisions. *Quarterly Journal of Economics* 133(4), 1885–1932.
- 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.
- Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *Quarterly Journal of Economics* 126(4), 1593–1660.
- Cooper, S. (2023). Prosecutor tradeoffs and race: Evidence from a circuit split.
- Didwania, S. (2025). Charging leniency and federal sentences. *Journal of Legal Studies*, Forthcoming.
- Didwania, S. H. (2020). Mandatory minimum entrenchment and the controlled substances act. *Ohio State Journal of Criminal Law* 18.
- Fischman, J. B. and M. M. Schanzenbach (2012). Racial disparities under the federal

sentencing guidelines: The role of judicial discretion and mandatory minimums. *Journal of Empirical Legal Studies* 9(4), 729–764.

Frandsen, B. R. (2017). *Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete*. Bingley, UK: Emerald Publishing Limited.

Gelman, A. and G. Imbens (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business and Economic Statistics* 37(3), 447–456.

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

Kleven, H. and M. Waseem (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *Quarterly Journal of Economics* 128, 669–723.

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

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

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy* 2(3), 180–212.

Shaffer, H. (2023). Prosecutors, race, and the criminal pipeline. *The University of Chicago Law Review* 90(7), 1889–1965.

Shaffer, H. and E. Harrington (2017). Brokers of bias in the criminal justice system: Do prosecutors compound or attenuate racial disparities inherited at arrest? <https://drive.google.com/file/d/1Y4r3yuYPX6cqsFtYh0MCKMY7FyC3QB80/view>.

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

The Pew Charitable Trusts (2015, August). Federal drug sentencing laws bring high cost, low return.

Tuttle, C. (2025a). Online appendix for ‘racial disparities in federal sentencing: Evidence from drug mandatory minimums’. [https://codytuttle.github.io/tuttle\\_mandatory\\_minimums\\_online\\_appendix.pdf](https://codytuttle.github.io/tuttle_mandatory_minimums_online_appendix.pdf).

Tuttle, C. (2025b). Racial disparities in federal sentencing: Evidence from drug mandatory minimums. [https://codytuttle.github.io/tuttle\\_mandatory\\_minimums.pdf](https://codytuttle.github.io/tuttle_mandatory_minimums.pdf).

United States Government Publishing Office (2023, September 22). National archives and records administration: Records schedules; availability and request for comments.

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

Yuan, A. Y. and S. Cooper (2022). Prosecutorial incentives and outcome disparities.

# Appendix

## A Additional tables

**Table A.1:** Drug data - summary statistics

	(1) All Cases	(2) Black	(3) Hispanic	(4) White
<i>Drug type:</i>				
cocaine	0.264	0.253	0.341	0.113
crack	0.181	0.461	0.0395	0.0453
heroin	0.142	0.144	0.164	0.0879
marijuana	0.172	0.0459	0.283	0.134
meth	0.241	0.0952	0.172	0.620
<i>Defendant characteristics:</i>				
criminal history points	4.112	5.698	2.394	5.330
female	0.124	0.0706	0.117	0.221
high school	0.569	0.661	0.423	0.739
age	35.43	35.91	34.04	37.68
<i>Outcomes:</i>				
gun involved	0.101	0.144	0.0595	0.122
trial	0.0295	0.0452	0.0233	0.0183
# of drug types charged	1.306	1.508	1.178	1.262
percent weight	0.736	0.669	0.753	0.805
conspiracy	0.479	0.501	0.466	0.474
MM imposed	0.219	0.271	0.200	0.177
safety valve	0.314	0.150	0.465	0.245
sentence length	72.59	93.85	58.84	68.82
Observations	42437	14122	19373	8942

*Notes:* Figures represent means. Sample includes all cases with weights between 20% and 180% of the threshold weight. All variables are binary except for criminal history points, age, percent weight, and sentence length. Sentence length is measured in months.

**Table A.2:** MM effects for all cases together

	(1) sent length	(2) resid sent 1	(3) resid sent 2	(4) resid sent 3
discontinuity	10.63	11.08	10.10	9.65
left side 95% CI	[70.36, 75.91]	[-0.61, 4.00]	[1.17, 5.58]	[1.62, 5.67]
right side 95% CI	[81.61, 86.23]	[10.96, 14.98]	[11.78, 15.61]	[11.77, 15.24]
selection gap	17.01	12.40	11.24	7.05
fit value	70%	70%	70%	70%
baseline controls	no	yes	yes	yes
fixed effects	no	no	yes	yes
pros. decision controls	no	no	no	yes
N of obs fit on	33691	33691	33691	33691

*Notes:* This table gives the regression discontinuity results for the full sample of cases. Column 1 presents the simple linear case fitted at 70% with no controls. Columns 2, 3, and 4 then add various controls (through residualization), still fitting at 70%. Baseline controls include criminal history points, drug type, and racial group, defendant age, defendant sex, and defendant education. Fixed effects are at the district and year level. Prosecutor decision controls include whether a gun was used in the offense, the number of drug types charged in the case, and whether the case went to trial.

**Table A.3:** MM effects with varying fits

	(1) sent length	(2) sent length	(3) sent length	(4) sent length
discontinuity	9.502	10.98	9.832	6.676
left side 95% CI	[71.36, 77.15]	[70.63, 77.22]	[69.34, 76.24]	[74.68, 79.39]
right side 95% CI	[81.61, 86.23]	[81.61, 86.23]	[81.61, 86.23]	[81.61, 86.23]
selection gap	17.01	17.01	17.01	17.01
fit value	60%-80%	64%	62%	79%
N of obs fit on	-	32177	31706	35067

*Notes:* This table repeats the simple linear discontinuity displayed in Table 2, but now uses various fitting schemes. Column 1 averages all fits between 60% and 80%, while columns 2-4 give the median, upper bound, and lower bound fits, respectively. Confidence intervals are determined using the standard error of the expected prediction.

**Table A.4:** MM effects without extrapolation

	(1)	(2)	(3)	(4)	(5)	(6)
	sent.	sent.	sent.	sent.	sent.	sent.
discontinuity	40.67 (2.21)	25.23 (1.94)	32.15 (1.80)	21.13 (1.30)	22.40 (1.31)	16.90 (1.15)
bandwidth	10%	10%	30%	30%	80%	80%
baseline controls	no	yes	no	yes	no	yes
N of obs fit on	42414	42414	42414	42414	42414	42414

*Notes:* This table presents traditional regression discontinuity estimates for the 10-year MM eligibility threshold. Across each specification the dependent variable is sentence lengths in months and the running variable is drug weight normalized to a percent of the threshold weight. Baseline controls include criminal history points, drug type, and racial group, defendant age, defendant sex, and defendant education.

**Table A.5:** Main effects by criminal history and safety valve

	(1) no prior history	(2) low history	(3) no safety valve	(4) yes safety valve
discontinuity	6.708	10.46	16.11	5.397
left side 95% CI	[1.41, 7.12]	[-2.16, 3.53]	[-3.79, 2.88]	[0.84, 3.43]
right side 95% CI	[8.11, 14.16]	[8.38, 14.25]	[13.01, 18.69]	[6.30, 9.04]
bunch effect	9.830	11.21	11.93	-5.984
fitted mean at 99%	4.26	0.68	2.14	2.61
fit value	70%	70%	70%	70%
N of obs fit on	3395	6828	16314	7197

*Notes:* The dependent variable is sentence length residualized against drug type and race. All specifications are discontinuities based on local linear fits. 95% confidence intervals are presented for each fit regression on either side of the cutoff. These are calculated using the standard error of the predicted expected value, and significance is determined as no overlaps between these two intervals. The causal effect is the regression discontinuity between the two extrapolated fits at the 10-year MM cutoff. Evidence of manipulation is seen in the difference between the right-hand regression fit and the actual sentence length at the threshold weight. Fit value represents the cutoff for where extrapolation begins. In this case, all specifications have left-hand regressions fit on cases with weights between 20% and 70% of the threshold weight, with extrapolation occurring from 71% up to the cutoff.

**Table A.6:** MM effects with varying right-hand side fits

	(1) sent length	(2) sent length	(3) sent length	(4) sent length	(5) sent length
discontinuity	11.05 (3.773)	10.74 (4.037)	11.27 (4.508)	12.32 (5.849)	13.44 (7.912)
$\delta_R$ value	110%	120%	130%	140%	150%
N of obs fit on	32818	30949	29419	28095	27152

*Notes:* This table repeats the simple linear discontinuity displayed in Table 2, but now varies  $\delta_R$ , the right-hand side fitted value. All regressions include controls for drug type, criminal history, education, and citizenship. Standard errors are clustered at the district level.

**Table A.7:** High vs low bunching districts

	(1) sent length	(2) resid sent 1	(3) resid sent 2	(4) resid sent 3
<i>Panel A: <math>\leq 3\%</math> bunching</i>				
discontinuity	9.674	9.387	9.426	9.465
N of obs fit on	25759	25759	25759	4812
<i>Panel B: <math>\leq 4\%</math> bunching</i>				
discontinuity	11.06	11.31	10.82	10.92
N of obs fit on	33390	33390	33390	33390
<i>Panel C: <math>\leq 5\%</math> bunching</i>				
discontinuity	10.98	11.34	10.53	10.63
N of obs fit on	37140	37140	37140	4812
<i>Panel D: <math>\geq 5\%</math> bunching</i>				
discontinuity	9.467	7.862	6.505	5.734
N of obs fit on	12784	12784	12784	12784
fit value	70%	70%	70%	70%
baseline controls	no	yes	yes	yes
fixed effects	no	no	yes	yes
pros. decision controls	no	no	no	yes

*Notes:* This table presents the main discontinuity effects for sub-samples of districts based on the proportion of cases at the bunching point. Panel A, B, and C give districts with low amounts of bunching while Panel D gives districts where more than 5% of all cases were at the bunching point. CI's are omitted for space, but are similar to mainline results. All extrapolation is fitting between 20% and 70% for the left side of the threshold and 105% and 180% on the right side.

**Table A.8:** Discontinuity analysis with multiple drug weight controls

	(1) All cases	(2) Black	(3) Hispanic	(4) White
discontinuity	9.520	11.86	10.18	11.02
left side 95% CI	[1.63, 5.64]	[2.60, 11.07]	[0.54, 4.97]	[-6.01, 2.01]
right side 95% CI	[11.65, 15.11]	[14.72, 23.23]	[10.89, 15.16]	[6.09, 12.42]
selection gap	5.573	3.665	3.856	-2.249
fit value	70%	70%	70%	70%
N of obs on fit	33691	11205	15296	7190

*Notes:* The dependant variable is residual sentence length. All specifications include controls for defendant characteristics, time and district fixed effects, and the new other drug type weight controls. These are weight as a percent of the 10-year mandatory minimum threshold for up to 4 other drug types other than the primary type. Each specification is fit at 70%. Selection gap gives the difference between the right fitted regression and the actual sentence length at the threshold.

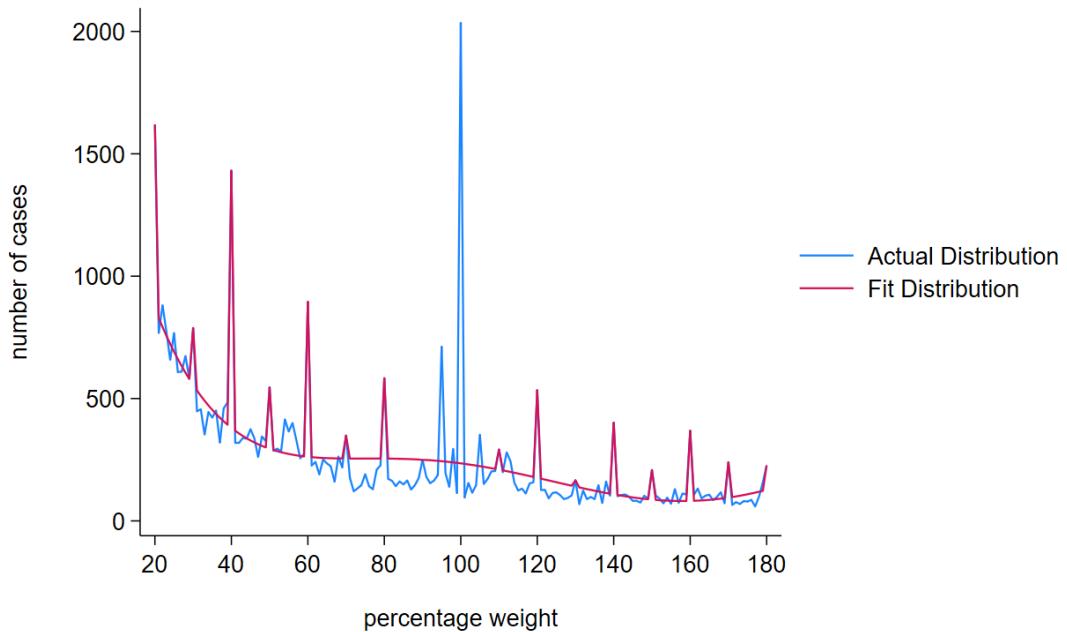
**Table A.9:** Discontinuity analysis with weight range observations included

	(1) sentence length	(2) resid sent 1	(3) resid sent 2	(4) resid sent 3
<i>Panel A: Minimum of weight range</i>				
discontinuity	9.467	13.00	9.192	11.02
left side 95% CI	[71.70, 76.97]	[-4.43, 0.11]	[-5.05, -0.84]	[-1.84, 2.19]
right side 95% CI	[81.55, 86.20]	[8.86, 12.96]	[4.37, 8.09]	[9.52, 13.03]
	{33911}	{33911}	{33911}	{33911}
selection gap	15.20	7.156	12.38	5.319
<i>Panel B: Median of weight range</i>				
discontinuity	13.69	11.68	10.91	11.09
left side 95% CI	[68.19, 73.37]	[-1.40, 3.077]	[-5.90, -1.76]	[-0.67, 3.26]
right side 95% CI	[82.19, 86.81]	[10.56, 14.64]	[5.15, 8.86]	[10.71, 14.21]
	{33546}	{33546}	{33546}	{33546}
selection gap	16.02	14.79	5.372	8.309
<i>Panel C: Maximum of weight range</i>				
discontinuity	13.22	16.93	9.799	12.06
left side 95% CI	[67.44, 72.39]	[-6.69, -2.42]	[-6.32, -2.40]	[-1.78, 1.94]
right side 95% CI	[80.86, 85.51]	[10.37, 14.47]	[3.52, 7.25]	[10.44, 13.97]
	{33546}	{33546}	{33546}	{33546}
selection gap	11.37	6.788	6.706	3.471
fit value	70%	70%	70%	70%
baseline controls	no	yes	yes	yes
additional controls	no	no	yes	yes
fixed effects	no	no	no	yes

*Notes:* This table gives the regression discontinuity results for the full sample of cases including cases with imprecise weight measures. For each measure, I use the precise weight for cases where it is available and then vary how the range measure is considered. In Panel A, I use the minimum value of the range as the drug weight measure. In Panel B, I use the median of the drug weight range. And in Panel C, I use the maximum weight in the drug range. The four specifications use the same control schemes as used in the main analysis. All regressions are fit at 70% of the threshold weight and use a linear fit. Sample size is given in curly braces.

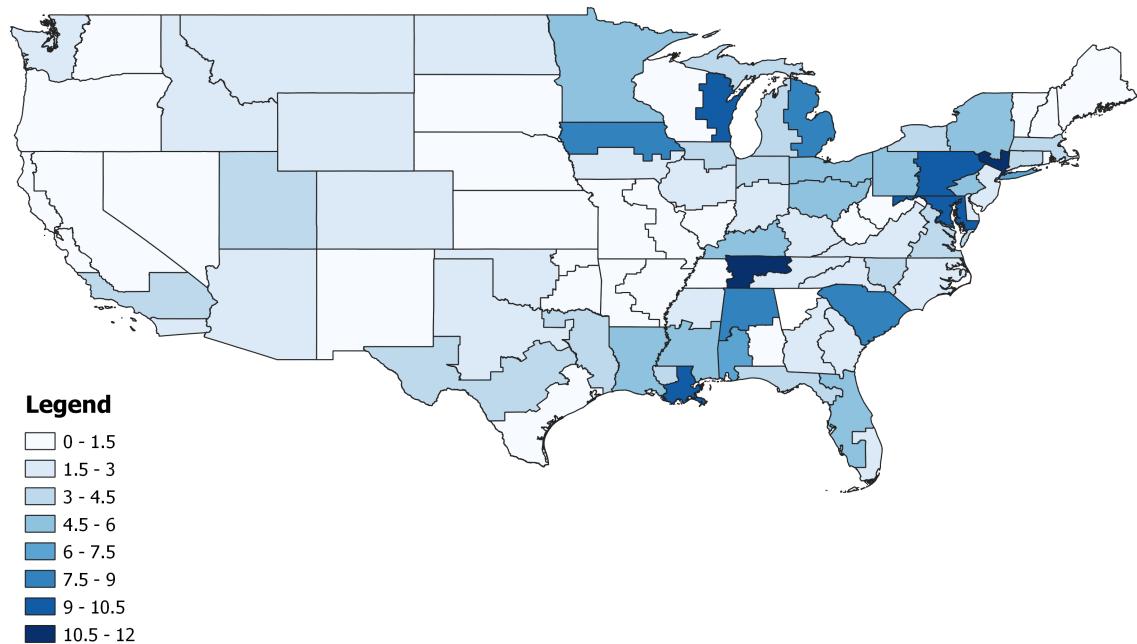
## B Additional figures

**Figure A.1:** Distribution of cases with fit



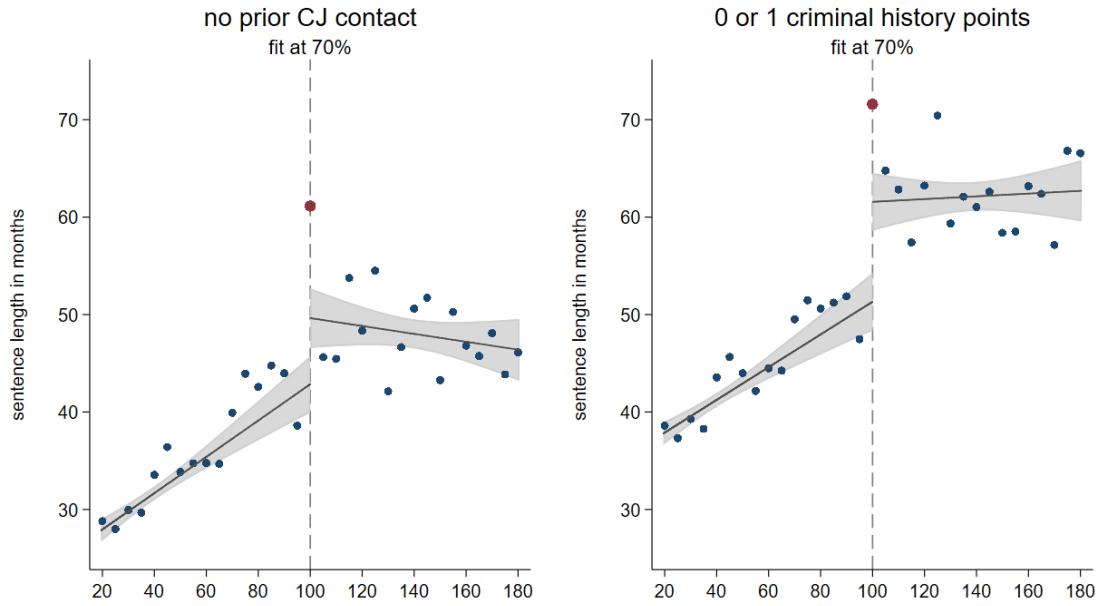
*Notes:* This figure considers which cases are being manipulated to the bunching point. I plot the distribution of cases and then fit a 5<sup>th</sup> order polynomial controlling for internal bunching that occurs at round points (I control for every 10% value). Missing mass is identified as areas where cases are below the fit polynomial - this appears to be primarily from weights at 70% to 95%.

**Figure A.2:** Percent of cases bunched by district



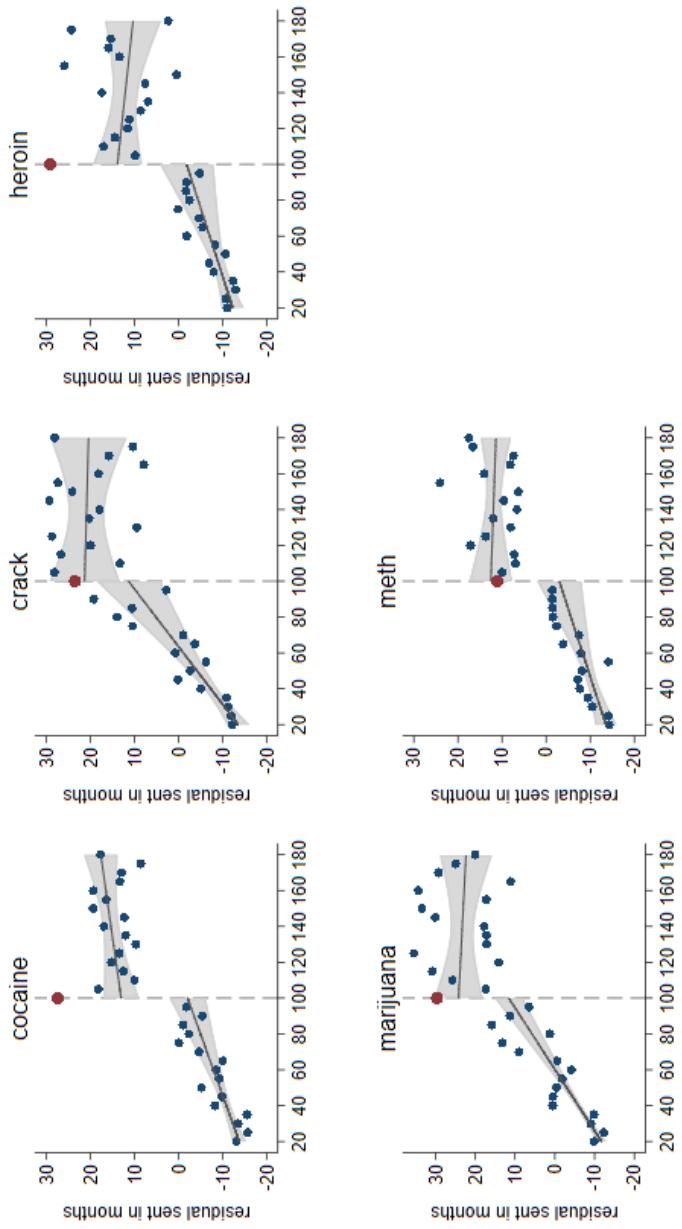
*Notes:* This figure shows variation in bunching propensity across district. It presents the percent of cases bunched for all cases before December of 2018 across each district.

**Figure A.3:** Main results by criminal history group



*Notes:* These are the main regression discontinuity results for cases in the first two criminal history categories, generally considered low-history defendants. Group 1 includes only defendants who have had no previous encounters with the criminal justice system, including events that would lead to zero criminal history points, such as arrest. Group 2 includes individuals who have no points but have had some encounters with the justice system, and individuals with one point. Both discontinuities are fit using the 70% cutoff for extrapolation and use linear fits.

**Figure A.4:** Main results by drug type



*Notes:* This figure illustrates the regression discontinuity design controlling for race and criminal history points, fit at 70% for each drug type. Data at the bunching point is larger and with a different color simply to emphasize differences in selection patterns between drug types.