

# How Much do Mandatory Minimums Matter?\*

Spencer Cooper

July 7, 2025

## Abstract

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.

**Address:** 365 Fairfield Way, U-1063, Storrs, CT, 06269.

**Email:** spencer.cooper@uconn.edu. **Phone:** 480-304-2767

**Thanks:** I thank Evan Taylor, Katherine Barnes, Tiemen Wouterson, Daniel Herbst, Hidehiko Ichimura, Jason Kreag, Stephen Ross, David Simon, Kevin Schnepel, 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 (2023) finds Blacks 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 (2023) 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.

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 distributions at the threshold captures the effect of MM eligibility absent manipulation.

I then 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 MM causally increases sentence lengths by about 10 months (or about 14% above the mean). This effect is consistent across defendant race, criminal history level, and drug type. The estimate is smaller, but still present for first-time offenders, suggesting that policies such as the safety valve provision may not adequately shield low-level offenders from MM effects. Because most of these cases are resolved through a plea bargain, the mechanism for this effect is likely increased bargaining power to the prosecutor.

Beyond estimating the average effect, I also find significant evidence of selection. The large gap between fitted expected 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. Interestingly, these selection patterns exist only for minority defendants. Even when conditioning on the most severe cases, I find no evidence of manipulation for white defendants, but significant distortions 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 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 2023; 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 (Yuan and Cooper 2022; Shaffer and Harrington 2017; Shaffer 2023). 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 2023). This paper uses discontinuity techniques similar to Diamond and Persson (2017), leniency measures similar to Goncalves and Mello (2021), and exploits the same bunching point as seen in Tuttle (2023).

Bjerk (2017b) is the paper most closely related to this work. It provides a descriptive analysis of the sentence length impacts of being convicted at particular charging weights for fiscal years 2011 and 2012. The main analysis shows that MM eligibility rates are similar across drug types, the safety valve provision effectively prevents low-level offenders from receiving large sentence length increases, and that effects for crack are minimal.<sup>1</sup> I contribute to this paper in four ways: (1) by identifying the causal impact of MM laws on sentence length accounting for selection, (2) by considering which types of defendants are manipulated to the threshold, (3) by assessing the degree to which low criminal history defendants are protected from MM effects, and (4) by giving a more comprehensive view of MM impacts on racial and drug type disparities. To my knowledge, this is the first paper to causally estimate the impacts of MM eligibility on sentence length.

---

<sup>1</sup>It's worth noting that in a brief appendix analysis, Tuttle (2023) finds that crack does indeed exhibit some sentence length effects if other years are included.

## 2 Background

### 2.1 Mandatory Minimum Sentencing of Federal Drug Cases

The main criteria for MM eligibility is drug quantity.<sup>2</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. Some cases 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 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 a another offender,<sup>3</sup> if the the offender is eligible for the safety valve

---

<sup>2</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.

<sup>3</sup>See USSG § 5K1.1

provision, or if a lower sentence is negotiated through plea bargain. Defense can apply for the safety valve provision if the defendant in question has a sufficiently low criminal history record. I directly test for MM effects for cases filed under the safety valve provision.

## **2.2 Drug Weight Manipulation**

Discrepancies between seized, charged, and sentencing drug weights can technically occur for several reasons. However, the primary source of drug weight manipulation occurs due to changes in relevant evidence. This may come by way of additional testimony or connecting a defendant to other traffickers or cases (Lynch 2016). The impetus for this additional evidence comes from prosecutors, who can decide whether to pursue additional evidence in building a case. If prosecutors can connect a defendant to other offenders or larger organizations, they may increase the available evidence of drugs to charge a person with.

One indication of this evidentiary channel is 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. This appears true in my data; I find that 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. It's also worth noting that conspiracy charge rates are similar across race, with White defendants having slightly more conspiracy charges than non-White defendants.

## **3 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>4</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 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 2023, 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 Online Appendix Table A.1.

Additional data are used for robustness checks and a supplementary analysis in the Online Appendix. These include the National Incident-Based Reporting System (NIBRS) from the FBI’s

---

<sup>4</sup>The other three substances subject to MM law are PCP, LSD, and fentanyl. These had too few observations for any meaningful analysis.

Uniform Crime Reporting program and a hand-collected data set of US Attorneys. Details of these data and their use are contained in the online appendix.

## 4 Empirical Strategy

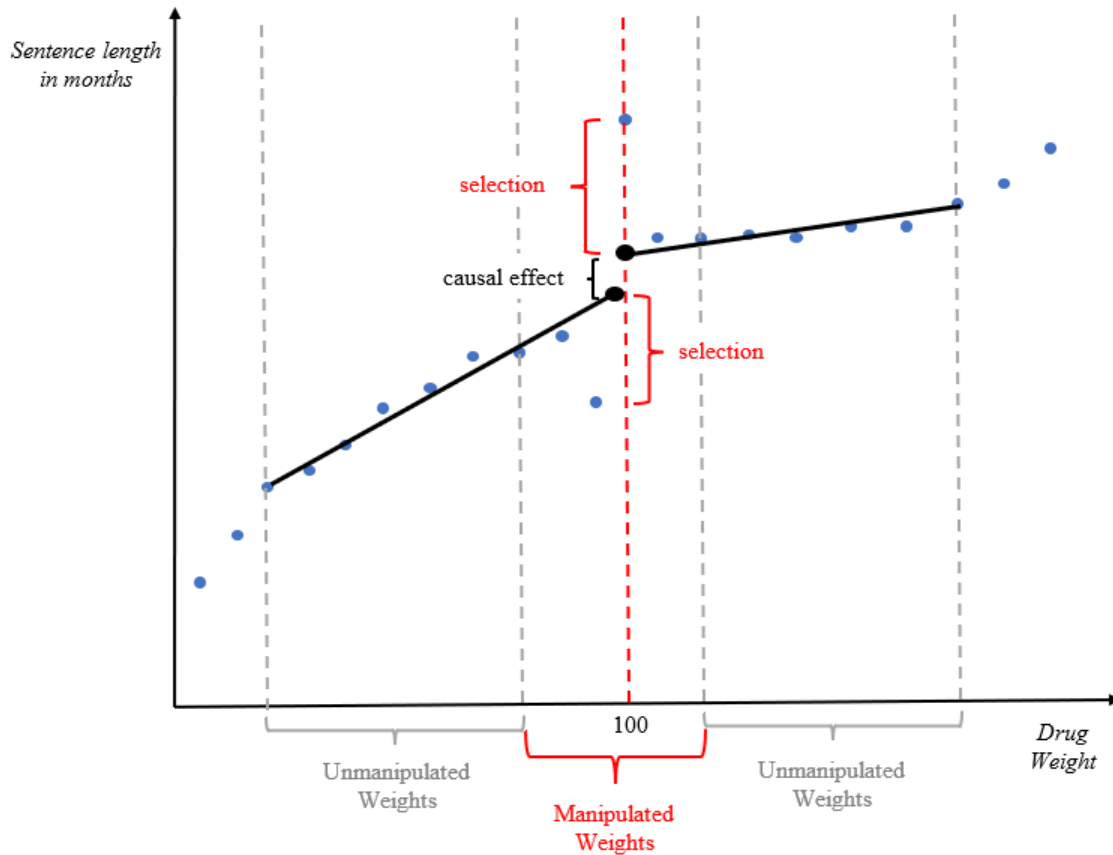
To accurately estimate the causal effects of MM sentencing, I need to know what the discontinuity at the MM threshold looks like without any manipulation. Thus, my strategy is to create fitted correlations between sentence lengths and charged drug weight over the range of unmanipulated drug weights. 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.

I also consider how these fitted distributions 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 1 illustrates the idea of this strategy. 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 1:** 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 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. Online 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. The discontinuity is considered statistically significant if the two confidence intervals do not overlap.

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, 2023), and the cost of gathering such evidence increases with the amount manipulated (Tuttle, 2022). 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; results are shown in Online Appendix Figure A.3. Third, Cooper (2023) finds that increased bunching due to policy changes occurs primarily near the threshold, suggesting manipu-

lation is concentrated among marginal cases. Fourth, 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 show that the RD estimates hold even in districts with minimal bunching, serving as a counterfactual where manipulation is less likely.

## 5 Results

### 5.1 Causal Effects

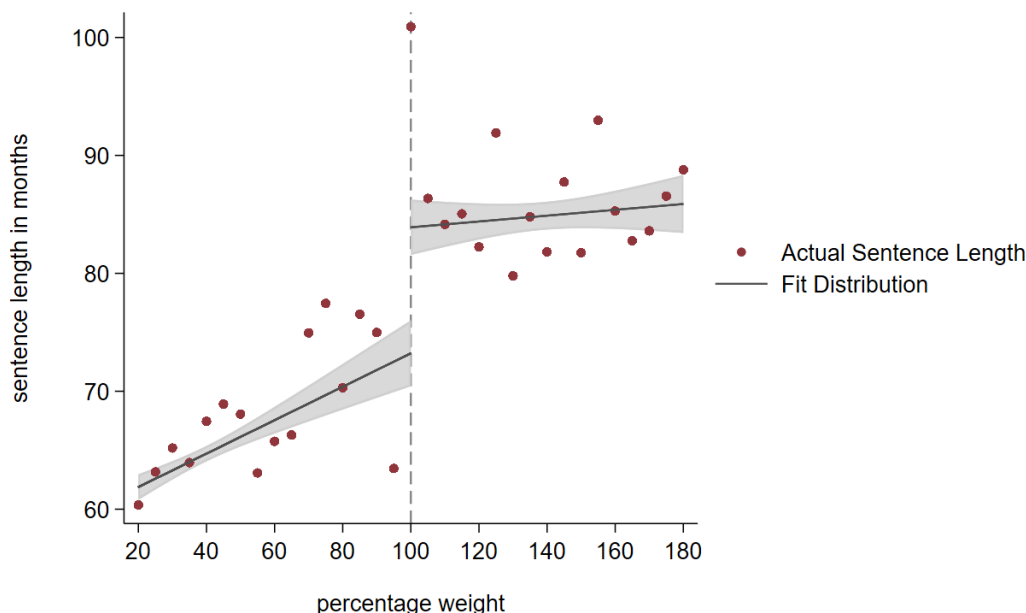
Figure 2 displays the aggregated discontinuity across all cases. MM eligibility is shown to increase sentence lengths by 10.63 months — a 13.9% rise over the predicted mean at the 99% threshold weight. The fitted values closely track actual sentence lengths within the manipulation range, except for cases 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>5</sup> The estimated 10-month discontinuity remains stable across all models. Full results are reported in Table A.2, with robustness checks using alternative fits shown in Table A.3 of the Online Appendix.

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.

---

<sup>5</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 went to trial.

**Figure 2:** 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.

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 Online Appendix Figure A.4.

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 Online Appendix Figure A.5.

**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, and significance is determined as no overlaps between these two intervals. 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%.

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.

## 5.2 Selection Patterns

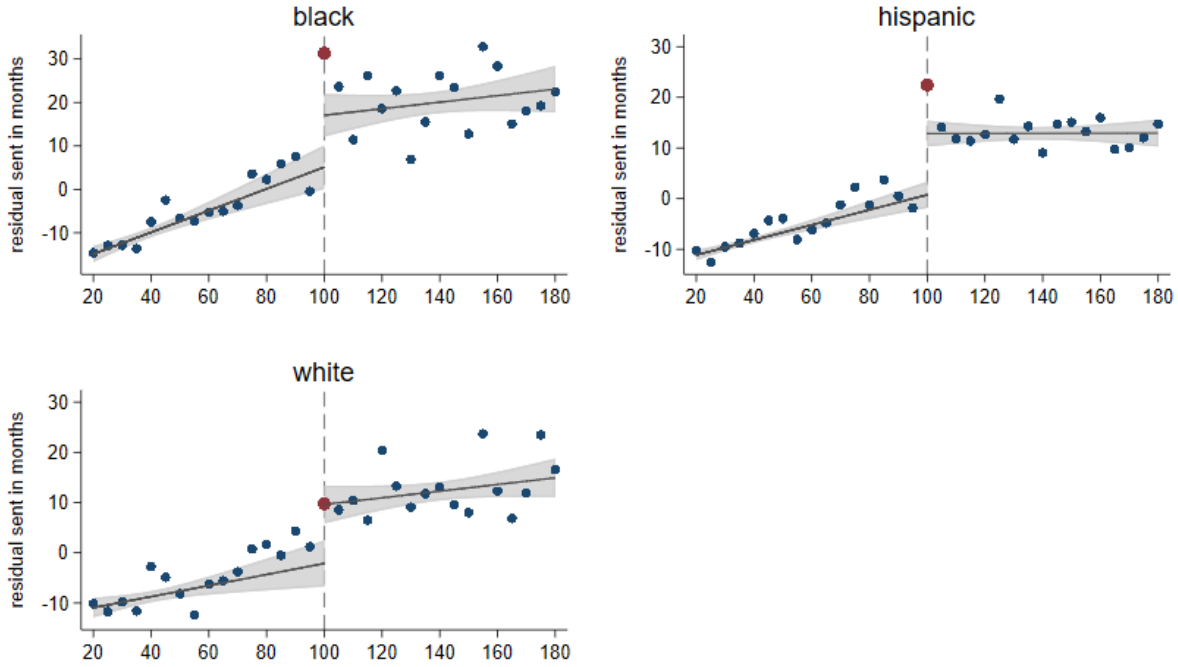
I now focus on patterns of selection by comparing cases within the manipulation region to their conditional expected sentence lengths. Although this analysis is descriptive rather than strictly causal, it sheds light on how prosecutors decide whom to manipulate. In particular, by examining deviations 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 2, 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 3 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 (2023), 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 con-

**Figure 3:** 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.

sider 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. I find evidence of significant selection across these variables with large distortions at the threshold weight. Figure A.6 shows these plots for only minority cases with cocaine or heroin as the primary drug type. Perhaps most interesting is that if you limit the sample to only include cases most likely to be manipulated, White defendants still exhibit no distortions, shown in Figure A.7. That is, even for the White defendants who are most likely to have their case manipulated, their sentence lengths do not deviate from predicted counterfactuals. This indicates that race is a primary driver of whether a case is manipulated or not.

### 5.3 Robustness

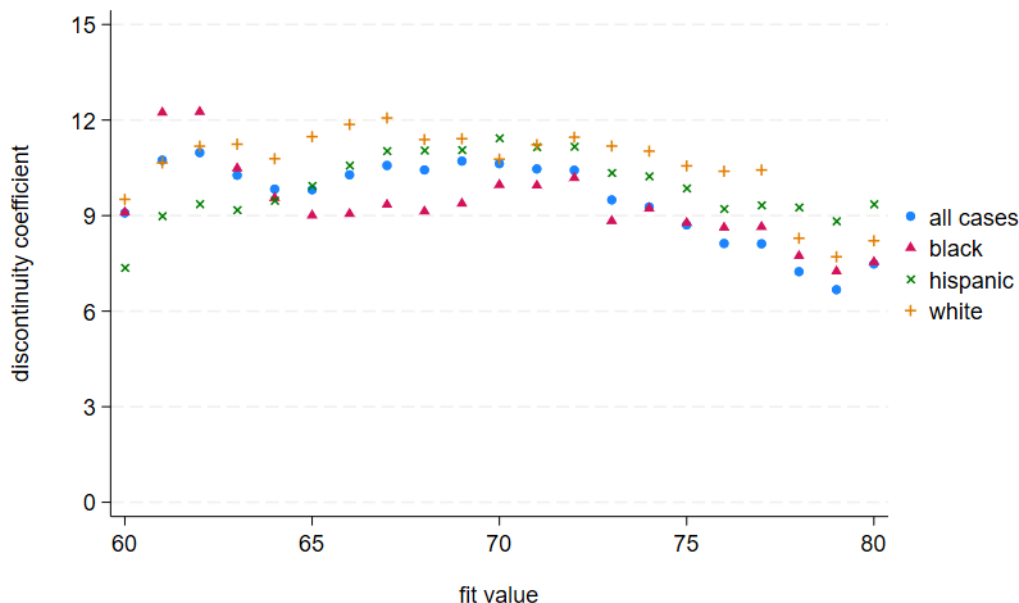
One of the primary concerns for the main results is that using the 70% cutoff for fitting over non-manipulated cases either creates bias (if 70% is too low) or doesn't capture all the variation in the data well (if 70% is too high). To check that results and magnitudes are not driven by one specific 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).

The heterogeneous effects presented in the main results are strongly consistent across different fit values. Groups with noisier data, like crack cases, tend to have higher variability of results by fit. Even still, these results tell a similar story. I present the RD coefficients across fit by race and for all cases in Figure 4. This graph shows the result for all 21 regressions for each group. Estimates are consistent across these specifications, suggesting results are not driven by specific cutoff choices.

I also show that the causal effects of MM eligibility are legitimate by comparing them 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 Online Appendix Table A.6.

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

**Figure 4:** 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 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.

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 is 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 Online Appendix Table A.7. 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 Online Appendix Table A.8. Across all three measures, the causal effects and selection impacts have similar magnitudes to main results and remain statistically significant.

## 6 Conclusion

In this paper, I show that mandatory minimum eligibility for federal drug cases affects sentence length in two ways: through a direct, statutory effect and by incentivizing prosecutors to manipulate cases. The average causal effect of eligibility is a 10.63 month increase in sentence length, which is a 13.9% increase over the counterfactual mean. This effect is consistent across case types including all race groups and low criminal history offenders. I find that actual sentence lengths near the threshold weight deviate significantly from fitted counterfactuals, giving evidence of prosecutor selection. Evidence of selection is only present for racial minority defendants. Even when limiting the sample to cases most likely to be manipulated, White defendants exhibit no selection patterns. This indicates that racial disparities in MM sentencing are driven by prosecutor charging decisions rather than by features of the MM statute.

## 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.
- Diamond, R. and P. Persson (2017). The long-term consequences of teacher discretion in grading of high-stakes tests. Nber working paper 22207, Stanford University.
- 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. (2022). Online appendix for ‘racial disparities in federal sentencing: Evidence from drug mandatory minimums’. [https://codytuttle.github.io/tuttle\\_mandatory\\_minimums\\_online\\_appendix\\_2022.pdf](https://codytuttle.github.io/tuttle_mandatory_minimums_online_appendix_2022.pdf).

Tuttle, C. (2023). 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.