Online Appendix for "Racial Disparities in Federal Sentencing: Evidence from Drug Mandatory Minimums"

Cody Tuttle

November 2022

Table of Contents

Appendix A. Additional Analyses

I. Tables

Table A1a. Summary Statistics for FL, NIBRS, and DEA Records

	Pre-2010	Post-2010	Observations
Panel A. Cocaine Felony Convictions in FL			
200-400g	0.00475	0.00432	214,573
	(0.0687)	(0.0656)	
28-200g	0.0405	0.0473	214,573
	0.197	(0.212)	
Missing drug weight	0.945	0.936	214,573
	(0.228)	(0.245)	
Black or Hispanic	0.771	0.789	214,573
	(0.420)	(0.408)	
Panel B. NIBRS Drug Seizures, Balanced Panel			
Weight (g)	10.31	7.85	191,667
	(45.61)	(45.75)	
280-290g	0.000361	0.000152	191,667
	(0.0190)	(0.0123)	
Black	0.737	0.745	191,667
	(0.440)	(0.436)	
Male	0.837	0.835	191,667
	(0.370)	(0.371)	
Panel C. DEA Drug Seizures			
Weight (g)	78.28	67.28	100,306
	(188.83)	(176.54)	
280-290g	0.00102	0.000428	100,306
	(0.0319)	(0.0207)	
Seized (vs. Purchased)	0.529	0.542	100,306
	(0.499)	(0.498)	
Price per gram (median)	47.36	56.18	37,280

Notes. The table above describes offenders found in the FL inmate database, the NIBRS drug seizure records, and the DEA drug exhibit data pre- and post-2010 (the DEA data actually describes the drugs themselves, not the offenders). The mean value of each variable is reported with standard deviations in parentheses. Observation counts are displayed separately for each variable. The statistics above are derived from the cleaned data in which the following cases are removed for NIBRS and DEA: cases with drug weights above 1000g. Weight is the weight of the drugs in grams recorded. 280-290g is a dummy variable equal to one when the weight is from 280-290g and zero when it is from 0-280g and 290-1000g, and missing when it is missing. The 200-400g and 28-200g variables follow the same logic. "Missing drug weight" is equal to one when the drug weight is missing. "Seized (vs. Purchased)" is equal to one if the DEA obtained the drug exhibit from a seizure versus an undercover purchase. The median price per gram is reported after removing outliers above the 95th percentile and below the 5th percentile.

Table A1b. Summary Statistics for EOUSA Prosecutor Case Files

	Pre-2010	Post-2010	Observations
Weight (g)	72.500	97.966	19,363
	(135.219)	(162.538)	
280-290g	0.004	0.082	19,363
	(0.062)	(0.274)	
280-290g, Missing = 0	0.002	0.026	49,342
	(0.040)	(0.158)	
50-60g	0.210	0.082	19,363
	(0.408)	(0.274)	
50-60g, Missing = 0	0.086	0.026	49,342
	(0.280)	(0.158)	
Missing drug weight	0.593	0.686	49,342
	(0.491)	(0.464)	
Only Federal Law Enforcement Involved	0.642	0.647	48,501
	(0.479)	(0.478)	
Any Federal Law Enforcement Involved	0.737	0.713	48,501
	(0.440)	(0.452)	
Lead Charge = Conspiracy	0.212	0.217	46,335
	(0.409)	(0.412)	

Notes. The table above describes defendants found in the EOUSA prosecutor case management data pre- and post-2010. The mean value of each variable is reported with standard deviations in parentheses. Observation counts are displayed separately for each variable since some fields in this data are missing much more often than others. The statistics above are derived from the cleaned data in which the following cases are removed: cases with drug weights above 1000g. Weight is the weight of the drugs in grams recorded in the case management system. 280-290g is a dummy variable equal to one when the weight is from 280-290g, zero when it is from 0-280g and 290-1000g, and missing when it is missing. "280-290g, Missing=0" is a dummy variable equal to "280-290g" but coded equal to zero when the weight field is missing. The 50-60g variables follow the same logic. "Missing drug weight" is equal to one when the drug weight is missing. "Only Federal Law Enforcement" is equal to one when the agency recorded as sending the case is strictly federal (i.e. DEA, FBI, or ATF) and equal to zero otherwise. "Any Federal" is equal to one if the agency sending the case has any federal involvement (i.e. "Joint DEA and state/local task force") and equal to zero otherwise. "Lead Charge = Conspiracy" is equal to one when the lead charge for the case is a drug conspiracy charge. The identical statistics in rows 3-10 of column 2 are not an error–in the EOUSA data, the number of 50-60g cases post-2010 is the exact same as the number of 280-290g cases post-2010.

Table A2. Exploring Bunching in Other Major Drug Types

Panel A. Racial differences in bunching in…	**Powder Cocaine**		**Heroin**		**Marijuana**		**Methamphetamine**		**Pooled**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	500-600g	5000-6000g	100-110g	1000-1100g	100kg-110kg	1000kg-1100kg	50-60g	500-600g	Lower	Upper
Black or Hispanic	0.00897**	0.00971***	0.00300	0.0135***	-0.00107	0.00183**	0.0330***	0.00769*	0.00446***	0.00602***
	(0.00420)	(0.00273)	(0.00684)	(0.00422)	(0.00146)	(0.000825)	(0.00937)	(0.00422)	(0.00146)	(0.00117)
Constant	0.0636***	0.0305***	0.0624***	0.0392***	0.0182***	-2.68e-05	0.138***	0.0631***	0.0362***	0.0200***
	(0.00738)	(0.00450)	(0.0114)	(0.00684)	(0.00229)	(0.00108)	(0.0169)	(0.00780)	(0.00234)	(0.00186)
Observations	27,926	51,658	9,309	18,453	67,792	83,774	7,550	17,282	128,555	171,166

Panel B. Racial Differences in Self-Reported Involvement								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Use Cocaine	Use + Sold Drugs	Use Heroin	Use + Sold Drugs	Use Marijuana	Use + Sold Drugs	Use Meth	Use + Sold Drugs
White	0.0576***	0.00282***	0.00290***	0.000981***	0.115***	0.000154	0.0344***	0.00202***
	(0.00153)	(0.000285)	(0.000585)	(0.000150)	(0.00215)	(0.000364)	(0.000782)	(0.000180)
	0.109***	0.00568***	0.0147***	0.00117***	0.347***	0.0119***	0.0205***	0.00158***
	(0.00124)	(0.000235)	(0.000500)	(0.000124)	(0.00178)	(0.000311)	(0.000563)	(0.000137)
Observations	763,466	762,181	763,565	762,275	763,297	762,012	763,622	762,322

Panel C. Racial Differences in Drug Seizures				
	(1)	(2)	(3)	(4)
	Cocaine Seized	Heroin Seized	Marijuana Seized	Meth Seized
White	9.692***	-0.127	354.8***	1.809*
	(2.335)	(0.892)	(21.54)	(1.077)
	28.54***	15.68***	-399.1***	19.51***
	(5.208)	(1.646)	(51.67)	(1.290)
Observations	140,370	74,861	1,791,564	162,287

Notes: Robust standard errors in parentheses. When possible, all specifications include fixed effects for federal district as well as controls for age and gender of the offender or respondent. For Panel A, the sample for each column is restricted to drug amounts between 0 and two times the threshold of interest (e.g., Panel A, column 1 is restricted to weights from 0-1000g; Panel A, column 2 is restricted to weights from 0-10000g). Following the main analysis, the range examined for bunching is a 10g range if the threshold is < 500g, a 100g range if the threshold is 500-1000g, a 1000g range if the threshold is above 1000g but below 10000g, and is 10% of the threshold for all amounts above 10000g. For Panel B, sample weights are applied. For Panel C, the sample for each column is restricted to drug amounts between 0 and two times the highest threshold of interest (e.g., Panel C, column 1 is restricted to weights from 0-10000g; Panel C, column 2 is restricted to weights from 0-2000g). These results differ from findings of a prior BJS working paper (BJS 2015) for two main reasons. First, the specifications in Panel A control for district fixed effects, this reduces the problems posed by round number bunching to some extent, since round number bunching varies across districts and defendant race varies across districts. Second, the specifications in Panel A are estimated on a wider range of cases, whereas the BJS working paper only examines the distribution narrowly around the threshold (e.g., Panel A, column 1 is estimated on any case from 0-1000g. The analogous specification in the BJS working paper restricts to 400-600g. This restriction will understate the level of bunching if cases below 400g are also "at risk" of bunching at 500g).

Difficulties in interpreting bunching in other drug types: The main analysis in the paper focuses on bunching in the distribution of crack-cocaine amounts. This is because crack-cocaine is the only drug for which the mandatory minimum threshold has changed since the introduction of the sentencing guidelines. The change in the threshold has two crucial benefits. First, the threshold was changed to 280g, which is a point with zero bunching before the change. This is important because all other thresholds are set at 50g, 500g, 1000g, etc., which are points that exhibit considerable round number bunching. In other words, there is bunching at these amounts even for drugs that do not have these amounts as relevant thresholds. The presence of round number bunching introduces noise that makes it difficult to determine if the observed bunching is due to prosecutor discretion re: sentence length concerns or due to the round number bunching phenomena. Second, a critical concern with interpreting a racial disparity in bunching at a threshold is that we do not know the underlying distribution of drug involvement for each drug type by race. In other words, suppose we don't observe a disparity in bunching at a threshold for marijuana. It's not clear if this is because marijuana involvement is higher for white offenders than for black and Hispanic offenders or if there is, in fact, no racial disparity conditional on drug involvement. The change in the threshold for crack-cocaine allows me to use the distribution of crack-cocaine amounts before the change to distinguish between those two possibilities, as outlined in the paper. This is not possible for the other drug types. Using data on self-reported drug use (Panel B) and drug seizures (Panel C), I show that this is a serious concern. White respondents/offenders do appear to be more involved with these other major drug types and thus the observed raw bunching likely understates any disparities in bunching that would exist conditional on drug involvement

Table A3. Result Robust to Other Drug Weight Sample Restrictions and Other Bunching Classifications

	Pr(280-290g Crack-Cocaine)					Pr(280-300g)	Pr(280-320g)	Pr(280-380g)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After 2010 x White	0.0127**	0.0123**	0.0123**	0.0605***	0.0597***	0.0288**	0.0156**	0.0144**	0.0103
	(0.0053)	(0.0051)	(0.0051)	(0.0092)	(0.0089)	(0.0124)	(0.0065)	(0.0070)	(0.0083)
After 2010 x Black or Hispanic	0.0343***	0.0328***	0.0327***	0.0832***	0.0820***	0.0715***	0.0346***	0.0352***	0.0376***
	(0.0021)	(0.0020)	(0.0020)	(0.0028)	(0.0027)	(0.0043)	(0.0022)	(0.0025)	(0.0029)
Constant	0.0018*	0.0017*	0.0016*	0.0026***	0.0032***	0.0023	0.0065***	0.0114***	0.0246***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0011)	(0.0027)	(0.0014)	(0.0019)	(0.0027)
P-value: W = BH	0.0001	0.0002	0.0002	0.0176	0.0166	0.0011	0.0053	0.0053	0.0020
Sample Restriction	0-2500g	0-25000g	No Restriction	0-1000g	0-1000g	50-1000g	0-1000g	0-1000g	0-1000g
Includes Weights Coded as a Range	No	No	No	Yes	Yes	No	No	No	No
Includes Weights Imputed from BOL	No	No	No	No	Yes	No	No	No	No
Observations	53,113	55,346	55,819	61,488	63,078	23,696	50,273	50,273	50,273

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Appendix E for notes about data construction. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." Columns 1-3 include outlier weights to varying extents. Column 4 reports results when the sample includes quantities coded as a range (in this analysis, the lower bound of the range is used). Column 5 reports results when the sample includes quantities coded as a range (using the lower bound of the range) and cases in which the weight is missing but a lower bound on the weight can be inferred from the base offense level (BOL). Column 6 excludes drug weights below 50g (i.e. excluding weights close to the 5-year mandatory minimum threshold pre- and post-2010). Columns 7-9 correspond to different definitions of what it means for a case to be "bunched" above the mandatory minimum threshold. For the main results, I define a result as "bunched" if it is in the narrow range of 280-290g. In columns 7-9, I use alternative ranges: 280-300g, 280-320g, and 280-380g.

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

Table A4. Result Robust to Various Sample Restrictions

	Pr(280-290g Crack-Cocaine)					
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0309***		0.0319***		0.0296***	
	(0.0021)		(0.0021)		(0.0022)	
After 2010 x White		0.0135**		0.0134**		0.0134**
		(0.0056)		(0.0057)		(0.0057)
After 2010 x Black or Hispanic		0.0321***		0.0330***		0.0307***
		(0.0022)		(0.0022)		(0.0023)
Constant	0.0050***	0.0031***	0.0061***	0.0031**	0.0062***	0.0031**
	(0.0004)	(0.0010)	(0.0006)	(0.0016)	(0.0007)	(0.0016)
P-value: W = BH	-	0.0018	-	0.0013	-	0.0047
Hispanic Offenders Excluded	Yes	Yes	No	No	Yes	Yes
Post-2006 Data Only	No	No	Yes	Yes	Yes	Yes
Observations	45,658	45,658	24,643	24,643	22,217	22,217

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Appendix E for notes about data construction. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "Post-2006 Data Only" is equal to "Yes" when the data is limited to cases brought to court from 2007-2015 (after the *Booker v. United States* Supreme Court case that made sentencing guidelines optional, excluding mandatory minimum guidelines). The row "Hispanic Offenders Excluded" is equal to "Yes" when Hispanic offenders are removed from the sample.
*** p<0.01, ** p<0.05, * p<0.1

Table A5. Result Robust to Controls and Alternative Std. Errors.

	Pr(280-290g Crack-Cocaine)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After 2010	0.0332***		0.0336***		0.0336***		0.0314***		0.0310***	
	(0.0067)		(0.0068)		(0.0066)		(0.0059)		(0.0058)	
After 2010 x White		0.0135**		0.0142**		0.0150**		0.0128**		0.0148**
		(0.0059)		(0.0060)		(0.0064)		(0.0062)		(0.0067)
After 2010 x Black or Hispanic		0.0344***		0.0349***		0.0347***		0.0325***		0.0319***
		(0.0070)		(0.0071)		(0.0069)		(0.0062)		(0.0062)
Constant	0.0049***	0.0031***	0.0041	0.0045	0.0057*	0.0070**	0.0045	0.0058	0.0050	0.0059*
	(0.0005)	(0.0011)	(0.0027)	(0.0028)	(0.0030)	(0.0030)	(0.0035)	(0.0035)	(0.0030)	(0.0031)
P-value: W = BH	-	0.0170	-	0.0179	-	0.0286	-	0.0289	-	0.0740
Offender Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Year Trend	No	No	No	No	No	No	Yes	Yes	Yes	Yes
State-specific Trends	No	No	No	No	No	No	No	No	Yes	Yes
Observations	50,273	50,273	50,273	50,273	50,273	50,273	50,273	50,273	50,273	50,273

Notes. Standard errors clustered at the state-level in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "Offender Controls" indicates if the following offender-level controls are included: criminal history points, age, sex, number of dependents, citizenship, number of current offense counts, whether a weapon was involved, and education. The rows "State Fixed Effects" and "Year Trend" indicate if the specification includes state fixed effects or a linear trend in year as controls. The row "State-specific Trends" indicates if the specification includes state-specific linear trends. In all cases, there is a sharp increase in the fraction of cases with 280-290g after 2010 and a racial disparity in that increase.
*** p<0.01, ** p<0.05, * p<0.1

Table A6. Result Robust to Probit, Logit, and Poisson Models.

	Probit			Logit			Poisson			OLS		
	280-290g		280-380g	280-290g		280-380g	280-290g		280-380g	280-290g		280-380g
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
After 2010 x W	0.6089***	0.6160***	0.1542	1.6969***	1.5489***	0.3618	1.2342***	1.2964***	0.1256	0.0135**	0.0282**	0.0103
	(0.1685)	(0.1882)	(0.1124)	(0.4618)	(0.4648)	(0.2595)	(0.4096)	(0.2793)	(0.5946)	(0.0056)	(0.0121)	(0.0083)
After 2010 x BH	0.8164***	0.8993***	0.3769***	2.0944***	2.1007***	0.8257***	2.1256***	2.1158***	0.8409	0.0344***	0.0681***	0.0376***
	(0.0352)	(0.0391)	(0.0243)	(0.0915)	(0.0924)	(0.0519)	(0.3661)	(0.2741)	(0.6294)	(0.0021)	(0.0042)	(0.0029)
Constant	-2.7399***	-2.4133***	-1.9671***	-5.7820***	-4.8323***	-3.6807***	3.4876***	2.5893***	3.5564***	0.0031***	0.0079***	0.0246***
	(0.1038)	(0.1148)	(0.0471)	(0.3167)	(0.3175)	(0.1132)	(0.3579)	(0.2249)	(0.3580)	(0.0010)	(0.0025)	(0.0027)
P-value: W = BH	0.2281	0.1405	0.0528	0.3986	0.2443	0.0796	0.0321	0.0035	0.3579	0.0004	0.0018	0.0020
Sample	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g
Observations	50,273	24,306	50,273	50,273	24,306	50,273	400	380	400	50,273	24,306	50,273

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x W" is equal to the coefficient on "After 2010 x BH," where "W" is the "White" dummy variable and "BH" is the "Black or Hispanic" dummy variables (abbreviated for table space). In general, columns 1-3 estimate probit models, columns 4-6 estimate logit models, columns 7-9 estimate Poisson models (on binned data), and columns 10-12 estimate OLS (or linear probability) models. Columns 1, 4, 7, and 10 estimate the change in bunching at 280-290g after 2010 for all cases from 0-1000g. Columns 2, 5, 8, and 11 limit the sample to cases from 50-1000g (following column 6 of Table A3). Columns 3, 6, 9, and 12 extend the "bunching" definition to 280-380g (following column 9 of Table A3). Although absolute increases in the probability of bunching are of interest because they correspond to the observed disparity in bunching, the estimates in columns 1-6 document that even relative to the pre-2010 disparity in charging at 280g, black and Hispanic offenders are more likely to be bunched than white offenders after 2010. These estimates are not statistically significant in columns 1-2 or 4-5 because the total number of white offenders in the 280-290g range is small and a small absolute increase can register as a large relative increase. Extending the range from 280-380g, as in columns 3 and 6, mitigates this small-N problem.

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

Table A7. Result Robust to Concerns about Selection Into/Out of Missing and Selection Into/Out of Other Drugs

	Pr(280-290g)			
	(1)	(2)	(3)	(4)
After 2010 x White	0.0583***	0.0247***	0.0005	0.0723***
	(0.0089)	(0.0060)	(0.0003)	(0.0033)
After 2010 x Black or Hispanic	0.0825***	0.0437***	0.0092***	0.2077***
	(0.0027)	(0.0021)	(0.0006)	(0.0031)
Constant	0.0034***	0.0025***	0.0004***	0.8680***
	(0.0009)	(0.0008)	(0.0001)	(0.0021)
P-value: W = BH	0.0089	0.0029	0.0000	0.0000
Drugs included	Crack-cocaine	Crack-cocaine	All	All
Dependent variable recoded to	Lower value of weight range	Upper value of weight range	Non-crack cases = 0	Non-crack cases = 1
Selection issue addressed	Into/out of missing weight	Into/out of missing weight	Into/out of other drugs	Into/out of other drugs
Observations	64,974	63,161	145,054	145,054

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Appendix E for notes about data construction. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "Drugs included" indicates the type of drugs included in the analysis. In columns 1 and 2, I focus on the crack-cocaine sample to analyze how including weights recorded as ranges affects the results. In columns 3 and 4, I focus on the sample of all drugs to analyze how movement of cases into or out of other drug types affects the results. The row "Dependent variable recoded to" indicates how the dependent variable is recoded in each analysis. In column 1, the dependent variable is recoded as 1 if the lower bound of the weight range is between 280-290g and recoded as 0 otherwise. In column 2, it is recoded as 1 if the upper bound of the range is between 280-290g and recoded as 0 otherwise. Results are also robust to recoding all missings as (In 280-290)=0 or recoding all missings as (In 280-290)=1. In column 3, the dependent variable is recoded as 0 if the case is not a crack-cocaine case, and in column 4, it is recoded as 1 if the case is not a crack-cocaine case. Columns 3 and 4 restrict the sample to weights between 0-1000g, but including outlier weights does not change the results (i.e. both columns would still indicate a similarly sized racial disparity that is statistically significant at the one percent level). Finally, the row "Selection issue addressed" indicates the type of selection issue being investigated in each column. In all columns, I find that the probability of being in the 280-290g range for crack-cocaine increases after 2010 and increases disproportionately for black and Hispanic offenders, regardless of selection into missing exact weights or other drug types.
*** p<0.01, ** p<0.05, * p<0.1

Table A8. Difference-in-Difference Bunching Identification

	Pr(280-290g)				Pr(50-60g)	
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0010*	-0.0002				
	(0.0006)	(0.0011)				
After 2010 x Crack-cocaine	0.0331***	0.0127**				
	(0.0021)	(0.0055)				
After 2010 x Crack-cocaine x Black or Hispanic		0.0214***				
		(0.0059)				
Crack-cocaine	-0.0024***	-0.0039***	-0.0041**	0.0089	0.0164***	0.0202*
	(0.0005)	(0.0011)	(0.0016)	(0.0059)	(0.0055)	(0.0122)
Crack-cocaine x Black or Hispanic			0.0021	0.0222***	0.0079	-0.0013
			(0.0018)	(0.0064)	(0.0058)	(0.0128)
Constant	0.0073***	0.0072***	0.0074***	0.0071***	0.0517***	0.0438***
	(0.0003)	(0.0006)	(0.0013)	(0.0027)	(0.0033)	(0.0065)
Drugs Included	All	All	Crack & Powder	Crack & Powder	Crack & Powder	Crack & Powder
Years Included	1999-2015	1999-2015	1999-2010	2011-2015	1999-2010	2011-2015
Observations	145,054	145,054	63,894	16,986	63,894	16,986

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. Columns 1-2 compare crack-cocaine cases to all other drug cases. Specifically, they estimate the change in the probability a case is recorded with 280-290g after 2010 both for crack-cocaine and for other drugs. Column 1 does this in general and column 2 does this by race. This amounts to a difference-in-difference (pre- vs. post-2010 and crack vs. non-crack) estimation of the bunching (as opposed to the pre- vs. post-2010 difference that is the focus of the paper). Columns 3-6 apply this same design to estimate the probability of being recorded with 280-290g and 50-60g before and after 2010. These columns compare crack to powder cocaine alone since powder cocaine is a drug that never has a 50g mandatory minimum threshold. Column 5 suggests that prior to 2010, black and Hispanic offenders are slightly more likely to be bunched at 50-60g specifically in crack-cocaine cases (34% of overall excess bunching for crack cases prior to 2010 and p-value = 0.18), but this difference disappears after 2010 (Column 6) when the threshold changes.

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

Table A9. Racial Disparity in Bunching at 280-290g without Departures Applied

	Pr(280-290g)		
	(1)	(2)	(3)
After 2010 x White	0.00367	0.00888*	0.00214
	(0.00326)	(0.00455)	(0.00268)
After 2010 x Black or Hispanic	0.0186***	0.0305***	0.0165***
	(0.00161)	(0.00202)	(0.00152)
Constant	0.00184**	0.00215***	0.00154**
	(0.000752)	(0.000812)	(0.000687)
P-value: W = BH	0.0000	0.0000	0.0000
Pr(280-290g) Recoded = 0 if...	-	-	-
...safety valve departure applied	No	Yes	Yes
...substantial assistance departure applied	Yes	No	Yes
Observations	50,273	50,273	50,273

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. Safety valve departures are uncommon for crack-cocaine offenses; they are received by 12% of all offenders and 9% after 2010. Substantial assistance departures are more common; they are received by 26% of all offenders and 22% after 2010. For cases in 280-290g range, 11% receive a safety valve departure after 2010 and 45% receive a substantial assistance departure. Slightly less than half of defendants in the 280-290g range receive at least one of the two departures. Column 1 recodes the dependent variable as zero if a substantial assistance departure is applied. Column 2 recodes the dependent variable as zero if a safety valve departure is applied. Column 3 recodes the dependent variable as zero if either departure is applied. These results indicate that there is a racial disparity in the probability a case is bunched at 280g **and** does not have a departure applied. In fact, the racial disparity is even more stark for this outcome. Note, these departures do reduce sentence length when applied, but to be eligible for these departures, defendants must cooperate with the government. This cooperation is also costly for defendants: it increases pressure to plea, it puts defendants and their families at risk of retaliation, and it potentially increases sentencing for other defendants. The extent to which the mandatory minimum fails to increase sentencing due to departures is directly related to the extent to which it increases these other costs. In *Hard Bargains*, Mona Lynch writes, "This practice can put defendants in a horrible dilemma, since providing assistance can be a deadly endeavor. [...] The problem for the defender's client was that he was from a neighborhood 'where if he cooperates, he's dead . . . or, even worse, his family's dead.'" For these reasons, I do not make a distinction between cases with or without a departure in the main analysis. *** p<0.01, ** p<0.05, * p<0.1

Table A10a. Missing Mass in the Distribution of Drug Amounts by Race

Panel A. Analysis of Changes in the 0-100g Range.					
	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 x White	0.0040	-0.1107***	0.0377**	-0.0036	0.0122
	(0.0185)	(0.0193)	(0.0155)	(0.0114)	(0.0146)
After 2010 x Black or Hispanic	0.0227***	-0.0717***	0.0342***	-0.0054*	-0.0102***
	(0.0040)	(0.0051)	(0.0042)	(0.0030)	(0.0039)
Constant	0.1945***	0.3202***	0.0965***	0.0679***	0.1017***
	(0.0069)	(0.0082)	(0.0052)	(0.0044)	(0.0053)
P-value: W = BH	0.3216	0.0505	0.8284	0.8769	0.1377
Observations	50,273	50,273	50,273	50,273	50,273
Panel B. Analysis of Changes in the 100-1000g Range.					
	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
	(6)	(7)	(8)	(9)	(10)
After 2010 x White	0.0005	0.0135**	0.0109	0.0061	0.0294***
	(0.0165)	(0.0056)	(0.0095)	(0.0067)	(0.0093)
After 2010 x Black or Hispanic	-0.0149***	0.0344***	0.0037	0.0021	0.0050**
	(0.0046)	(0.0021)	(0.0026)	(0.0018)	(0.0021)
Constant	0.1484***	0.0031***	0.0350***	0.0160***	0.0166***
	(0.0062)	(0.0010)	(0.0032)	(0.0022)	(0.0022)
P-value: W = BH	0.3705	0.0004	0.4638	0.5669	0.0103
Observations	50,273	50,273	50,273	50,273	50,273

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." See Table 3 for notes about the estimating equation. The only difference is that the specifications in this table include race interactions as described in the notes of Table 2.
*** p<0.01, ** p<0.05, * p<0.1

Table A10b. Racial Disparity in Shifting Above 290g vs. Above 280g After 2010

	Above 290g	Above 280g	Above 290g	Above 280g	Above 290g	Above 280g
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010 x White	-0.0183	-0.0051	-0.0252	-0.0093	-0.0306	-0.0153
	(0.0294)	(0.0308)	(0.0302)	(0.0315)	(0.0317)	(0.0331)
After 2010 x Black or Hispanic	-0.0087	0.0213**	-0.0010	0.0278***	0.0010	0.0292***
	(0.0079)	(0.0086)	(0.0079)	(0.0086)	(0.0084)	(0.0091)
Constant	0.1136***	0.1187***	0.1317***	0.1381***	0.1318***	0.1377***
	(0.0194)	(0.0198)	(0.0230)	(0.0234)	(0.0230)	(0.0234)
Trends Interacted with After2010	Quadratic	Quadratic	Linear	Linear	Linear	Linear
Post-2006 Only	No	No	Yes	Yes	Yes	Yes
Pre-2014 Only	No	No	No	No	Yes	Yes
P-value: W = BH	0.7534	0.4104	0.4395	0.2556	0.3358	0.1943
Observations	50,273	50,273	24,643	24,643	21,991	21,991

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." All specifications estimate trend breaks:

$$(ChargedX - Yg)_{it} = \alpha_0 + \beta_1(After2010 \times W)_{it} + \beta_2(After2010 \times BH)_{it} + \delta_1(After2010 \times W \times Trend)_{it} + \delta_2(After2010 \times BH \times Trend)_{it} + \gamma_1 BH + \phi_1(BH \times Trend) + Z_i + Trend_t + \epsilon_{it}$$

$Trend$ is either a linear trend or a quadratic trend centered at zero in 2011. The polynomial order of the time trend is indicated in the row "Trends Interacted with After2010." Columns (1), (3), and (5) estimate the change in the probability of being charged with an amount above 290g after 2010 by race. Columns (2), (4), and (6) estimate the change in the probability of being charged with an amount above 280g after 2010 by race. Columns (1)-(2) use the full time period and quadratic trends because the likelihood of being charged with an amount above 290g falls from 2000-2005 and begins rising in 2005. Columns (3)-(4) use the years 2007-2015 and linear trends because that period excludes the fall from 2000-2005. Columns (5)-(6) use the years 2007-2013 because the likelihood of being charged with an amount above 280g falls after *Alleyne v. US* in 2013. All specifications indicate that there is a racial disparity in being charged above 280g, in general. The disparity is noisy, but it is the same magnitude as the disparity in the relevant 280-290g range, and when I focus on that range (e.g. Table 2), the estimates are more precise.
*** p<0.01, ** p<0.05, * p<0.1

Table A10c. Missing Mass in the Distribution of Drug Amounts, Post-2007 Only & Trials Only

Panel A. Analysis of Changes in the 0-100g Range, Post-2007 Only					
	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
After 2010 (Actual Change)	0.0255***	-0.0703***	0.0318***	-0.0094***	-0.0103**
	(0.0043)	(0.0057)	(0.0045)	(0.0034)	(0.0043)
Constant	0.1074***	0.2925***	0.1134***	0.0750***	0.1263***
	(0.0025)	(0.0036)	(0.0025)	(0.0021)	(0.0027)
Predicted Change from Conceptual Model	Increase	Decrease	Increase	Decrease	Decrease
Observations	24,643	24,643	24,643	24,643	24,643
Panel B. Analysis of Changes in the 100-1000g Range, Post-2007 Only					
	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
After 2010 (Actual Change)	-0.0110**	0.0319***	0.0036	0.0024	0.0059***
	(0.0051)	(0.0021)	(0.0028)	(0.0019)	(0.0023)
Constant	0.1882***	0.0061***	0.0444***	0.0207***	0.0261***
	(0.0031)	(0.0006)	(0.0016)	(0.0011)	(0.0013)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	24,643	24,643	24,643	24,643	24,643
Panel C. Analysis of Changes in the 0-100g Range, Trials Only					
	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
After 2010 (Actual Change)	0.0280	-0.0722***	0.0212	-0.0007	-0.0596***
	(0.0191)	(0.0224)	(0.0177)	(0.0145)	(0.0134)
Constant	0.1185***	0.2742***	0.1000***	0.0765***	0.1151***
	(0.0070)	(0.0097)	(0.0065)	(0.0058)	(0.0069)
Predicted Change from Conceptual Model	Increase	Decrease	Increase	Decrease	Decrease
Observations	2,515	2,515	2,515	2,515	2,515
Panel D. Analysis of Changes in the 100-1000g Range, Trials Only					
	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
After 2010 (Actual Change)	-0.0234	0.0674***	0.0010	0.0252**	0.0131
	(0.0207)	(0.0129)	(0.0128)	(0.0118)	(0.0115)
Constant	0.1925***	0.0033***	0.0571***	0.0278***	0.0349***
	(0.0086)	(0.0012)	(0.0050)	(0.0036)	(0.0040)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	2,515	2,515	2,515	2,515	2,515

Notes. Robust standard errors in parentheses. See Table 1 for notes about data construction. Panels A-B restricts to cases sentenced from 2007-2015. Panels C-D restricts to cases that end in a jury trial. The predicted change from the conceptual model of prosecutor behavior in Section II.B is displayed in the row labeled "predicted change from conceptual model." See Table 3 for notes about the estimating equation.

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

Table A11. Relationship between 2012-2013 Bunching and 2012-2013 Appeals at the District Level or State Level

	Any Drug Appeal	Any Drug Appeal with Drug Amount Issue	Any Drug Appeal with Mand. Min. Issue	Any Drug Appeal, Reversed	Any Drug Appeal with Drug Amount Issue, Reversed	Any Drug Appeal with Mand. Min. Issue, Reversed	Any Drug Appeal, Reversed or Remanded	Any Drug Appeal with Drug Amount Issue, Reversed or Remanded	Any Drug Appeal with Mand. Min. Issue, Reversed or Remanded
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Unweighted regressions, district level									
Any Bunching in 2012-2013	0.118**	0.0571	0.244**	0.156	0.171**	0.0839	0.102	0.142	0.104
	(0.0577)	(0.0825)	(0.100)	(0.101)	(0.0696)	(0.0598)	(0.107)	(0.089)	(0.0688)
Constant	0.860***	0.791***	0.256***	0.279***	0.0465	0.0465	0.442***	0.163***	0.0698*
	(0.0534)	(0.0627)	(0.0673)	(0.0692)	(0.0325)	(0.0325)	(0.0766)	(0.0569)	(0.0393)
Panel B. Regressions weighted by total cases in district									
Any Bunching in 2012-2013	0.0943*	0.142	0.329***	0.0788	0.199**	0.152*	0.110	0.183	0.193**
	(0.0492)	(0.0956)	(.124)	(0.135)	(0.0863)	(0.0815)	(0.125)	(0.116)	(0.0964)
Constant	0.905***	0.765***	0.319***	0.422***	0.0395	0.0330	0.550***	0.177**	0.0683
	(0.0492)	(0.0871)	(0.0954)	(0.100)	(0.0359)	(0.0247)	(0.0979)	(0.0756)	(0.0426)
Observations	89	89	89	89	89	89	89	89	89
Panel C. Unweighted regressions, state level									
Any Bunching in 2012-2013	0.190**	0.190**	0.465***	0.279**	0.181*	0.0903	0.452***	0.305**	0.194**
	(0.0875)	(0.0875)	(0.126)	(0.134)	(0.107)	(0.0808)	(0.129)	(0.122)	(0.0940)
Constant	0.810***	0.810***	0.190**	0.238**	0.0952	0.0476	0.238**	0.143*	0.0476
	(0.0875)	(0.0875)	(0.0875)	(0.0949)	(0.0654)	(0.0474)	(0.0949)	(0.0779)	(0.0474)
Panel D. Regressions weighted by total cases in state									
Any Bunching in 2012-2013	0.146	0.146	0.419**	0.319*	0.244	0.157	0.421**	0.285	0.300**
	(0.0960)	(0.0960)	(0.189)	(0.188)	(0.179)	(0.110)	(0.180)	(0.197)	(0.124)
Constant	0.854***	0.854***	0.414**	0.401**	0.154	0.0497	0.401**	0.287*	0.0497
	(0.0960)	(0.0960)	(0.176)	(0.166)	(0.135)	(0.0509)	(0.166)	(0.163)	(0.0509)
Observations	50	50	50	50	50	50	50	50	50

Notes. Robust standard errors in parentheses. The USSC also provides data on cases appealed in the 12 circuit courts of appeal. This appeals dataset is separate from the main USSC dataset used throughout this paper. I focus on cases with appeal dates in fiscal year 2012 and 2013 since cases appealed before those fiscal years will not be subject to the FSA or to the new 2011 USSC Guidelines and cases appealed after fall under a new data structure that does not provide as much detail. I focus on bunching for cases sentenced from 2012-2013 to closely match the appealed cases under consideration. Although the appeals records are not linked to the main USSC case records, they do indicate whether the appeal raised an issue with the drug conviction and specifically, whether an issue was raised regarding the quantity of the drugs specified in the case and/or the application of the mandatory minimum sentence. Data after fiscal year 2013 do not provide this specific information about the appealed issues. Finally, the data include the outcome of the appeal, such as whether the issue was reversed (i.e. the appeals court finds the lower court decided the issue improperly) or remanded (i.e. the appeals court sends the case back to the lower court to review the issue again). In Panels A-B, each column is a regression of whether a district has any case which is: a drug appeal, a drug appeal with a drug amount issue, etc. on whether the district has any case that is bunched at 280-290g. Panel B weights those regressions by the total number of cases in the district. Results are similar when looking at number of appeals as a fraction of all sentenced drug cases in the district instead of the binary outcome of whether there was an appeal at all. Panels C-D have the same structure but are estimated at the state level instead of the district level. This provides additional evidence along with the results from *Alleyne* that bunching at 280g may be based on weak evidence or that the practice of bunching may be related to pursuing weak evidence in other drug cases.

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

Table A12. Bunching Analysis for Potential Mechanisms, Alternative Results

Panel A. Analysis of Bunching in State Convictions and in Drug Seizures				
	Pr(200-400g)	Pr(200-400g)	Pr(200-400g)	Pr(200-400g)
	(1)	(2)	(3)	(4)
After 2010	0.00358		0.0185	
	(0.00873)		(0.0444)	
After 2010 x White		0.0068		-0.0008
		(0.0116)		(0.0554)
After 2010 x Black or Hispanic		0.0017		0.0192
		(0.0095)		(0.0488)
Constant	0.103***	0.1018***	0.2132***	0.1615***
	(0.00616)	(0.0068)	(0.0297)	(0.0379)
Data Analyzed	FL Convictions	FL Convictions	NC Convictions	NC Convictions
Drugs Included	Cocaine, all types, Weight Only	Cocaine, all types, Weight Only	Cocaine, all types	Cocaine, all types
P-value: W = BH	-	0.6484	-	0.7270
Observations	12,194	12,194	843	843
Panel B. Analysis of Bunching in Drug Seizures and Final Sentencing				
	Pr(280-290g)	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)
	(6)	(7)	(8)	(9)
After 2010	-0.000174*		0.0323*	
	(9.20e-05)		(0.0167)	
After 2010 x White		0.0002		0.0021
		(0.0002)		(0.0554)
After 2010 x Black or Hispanic		-0.0003***		0.0339**
		(0.0001)		(0.0169)
Constant	0.000428***	0.0003***	0.145***	0.1667***
	(5.11e-05)	(0.0001)	(0.0124)	(0.0232)
Data Analyzed	NIBRS, Full Coverage States	NIBRS, Full Coverage States	USSC Sentencing, NC only	USSC Sentencing, NC only
Drugs Included	Crack-cocaine	Crack-cocaine	Cocaine, all types	Crack-cocaine
P-value: W = BH	-	0.0833	-	0.5656
Observations	207,043	207,043	4,267	4,267

Notes. Robust standard errors in parentheses. When possible, the specifications above use a sample of offenses with drug amounts between 0 grams and 1000 grams. Analyses of state-level drug convictions do not make this restriction since the state reports broad drug weight categories instead of specific amounts. When broad categories (200-400g) are analyzed, a linear trend in year is included. The row "P-value: W= BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." In Panel A: columns 1-2 show an analysis of reported drug amounts for state-level drug convictions in Florida that restricts to cases where some weight range is listed in the offense description, columns 3-4 show an analysis of state-level drug convictions in North Carolina (a state where only some offenses specify the type of drug involved). Columns 5-6 show an analysis of weights for seized drugs reported to the FBI through the National Incident Based Reporting System (limiting to states that have full coverage from 2012-2015 and have at least 90% coverage from 2008-2015), Finally, columns 7-8 show an analysis of weights from USSC sentencing data for federal convictions in NC using broad drug categories and all types of cocaine.
*** p<0.01, ** p<0.05, * p<0.1

Table A13. Variation in Bunching at 280-290g By Type of Agency Sending the Case

	280-290g	280-290g	280-290g	Weight (g)
	(1)	(2)	(3)	(4)
After 2010	0.0820***	0.0752***	0.0993***	26.10***
	(0.0179)	(0.0190)	(0.0130)	(5.677)
After 2010 × Any Federal	0.000730			
	(0.00133)			
After 2010 × Only Federal		0.000454		
		(0.00148)		
After 2010 × FBI			0.0157	52.96***
			(0.0198)	(11.30)
After 2010 × ATF			-0.0735***	-15.00**
			(0.0144)	(6.967)
After 2010 × State/local			-0.0241	-7.666
			(0.0230)	(11.38)
After 2010 × DEA & State/local			-0.0136	-3.996
			(0.0383)	(19.46)
After 2010 × Joint state/local			0.0144	7.215
			(0.0507)	(25.60)
After 2010 × ATF & State/local			-0.00897	-9.403
			(0.0388)	(13.19)
After 2010 × FBI & State/local			-0.0623	-17.34
			(0.0386)	(22.44)
Constant	0.00341***	0.00359***	0.00481***	77.81***
	(0.00120)	(0.00135)	(0.000877)	(1.525)
Observations	17,043	15,017	17,043	17,043

Notes. Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. Column 1 interacts the after 2010 dummy variable with a dummy variable equal to one when the agency recorded as sending the case involves a federal agency (i.e. DEA, ATF, FBI). This includes agencies recorded as a federal agency joint with a state/local task force. Column 2 interacts the after 2010 variable with a variable equal to one when the agency sending the case is strictly federal (i.e. not including any involvement from state/local authorities). Column 2 does not include "joint" investigations in the sample. Column 3 provides more detail by interacting the after 2010 dummy variable with dummy variables for the top agencies (with the DEA as the reference category). Most agencies have similar levels of bunching at 280-290g post-2010. Two agencies have considerably lower levels, but as column 4 shows, those agencies are involved with lower drug weight cases, in general.
*** p<0.01, ** p<0.05, * p<0.1

Table A14. Offender Drug-Holding Behavior by Race, After Fair Sentencing Act in 2010, Full Coverage States

	Weight	Pr(280-290g)	Weight	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-280g)	Pr(270-280g)	Pr(280-290g)	Pr(>290g)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After 2010 x White			-0.6018	0.0302***	-0.0210***	-0.0033**	-0.0058***	-0.0001	0.0001	-0.0002
			(0.5999)	(0.0041)	(0.0037)	(0.0016)	(0.0013)	(0.0000)	(0.0002)	(0.0007)
After 2010 x Black			-2.8015***	0.0403***	-0.0172***	-0.0064***	-0.0143***	-0.0001***	-0.0002**	-0.0020***
			(0.2504)	(0.0027)	(0.0025)	(0.0011)	(0.0009)	(0.0000)	(0.0001)	(0.0003)
Black	2.503***	9.21e-05	3.0414***	-0.1125***	0.0825***	0.0137***	0.0148***	0.0001	0.0002	0.0013***
	(0.260)	(0.000102)	(0.2885)	(0.0025)	(0.0023)	(0.0010)	(0.0009)	(0.0001)	(0.0001)	(0.0004)
Constant	10.01***	0.000454***	9.7586***	0.7503***	0.1856***	0.0310***	0.0284***	0.0002**	0.0004***	0.0043***
	(0.426)	(0.000152)	(0.4417)	(0.0040)	(0.0036)	(0.0016)	(0.0014)	(0.0001)	(0.0002)	(0.0006)
Observations	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043
P-value: W = B	-	-	0.0007	0.0408	0.3969	0.1075	0.0000	0.3308	0.1266	0.0205

Notes. Robust standard errors in parentheses. This analysis uses the weights of seized drugs reported to the FBI through the National Incident Based Reporting System. Ethnicity is not consistently recorded in NIBRS over this time period. As such, I refer to offenders as black or white, omitting the Hispanic label used in previous analyses. Columns 1-2 show the relationship between race of offender and drug weight seized, in general. Column 3 shows how the weight of an offender's seized drugs changes by race after 2010. Columns 4-10 show how the probability an offender's seized drugs are in a certain bin changes by race after 2010. All specifications include state fixed effects and controls for age and sex. The row "P-value: W= B" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black." The sample is limited to states that have full coverage from 2012-2015 and have at least 90% coverage from 2008-2015.

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

Table A15. Relationship between Bunching in EOUSA and Imputed Defendant Race

	280-290g, Missing = 0	280-290g	280-290g, Missing = 0	280-290g	280-290g, Missing = 0
	(1)	(2)	(3)	(4)	(5)
After 2010	0.0241***	-0.0318	-0.0153**	-0.00536	-0.00511
	(0.00180)	(0.0196)	(0.00654)	(0.0229)	(0.00826)
After 2010 × % Black or Hispanic		0.123***	0.0457***	0.0793***	0.0303***
(for Cases Sentenced in District-Month)		(0.0295)	(0.01000)	(0.0282)	(0.00984)
Constant	0.00159***	-0.00193	-0.00111	-0.00202	-0.000842
	(0.000195)	(0.00319)	(0.00130)	(0.00633)	(0.00259)
Prosecutor FEs	NO	NO	NO	YES	YES
Observations	49,342	13,384	32,751	13,384	32,751

Notes. Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. Column 1 displays the main bunching result using a dependent variable that is equal to one when the drug weight in the case is between 280-290g and is equal to zero if it is not in that range. Importantly, "280-290g, Missing=0" is also coded as zero if the drug weight field is missing. This is especially relevant for cross-district analyses because weight missingness varies substantially across districts. Robustness to alternative ways of dealing with missing values is further explored in Appendix E. Coefficients are estimated from the following regression for column 1:

$$(Charged\ 280-290g,\ Missing=0)_{it} = \alpha_0 + \beta_1 After2010_{it} + \epsilon_{it}$$

Columns 2-5 interact the after 2010 dummy variable with a probabilistic estimate of defendant race (race is not available in the EOUSA files). To impute defendant race, I match EOUSA information about sentence year-month to USSC information about the racial composition of sentences in each district year-month. I code "% Black or Hispanic" equal to the fraction of offenders in a district and sentenced in a year-month who are black or Hispanic. In columns 4-5, I include prosecutor fixed effects. Specifications with the race and after 2010 interactions also include a variable equal to % black and Hispanic offenders in the district-year-month. The number of observations falls because not all cases that enter EOUSA end in a sentence or have sentence information recorded. Coefficients are estimated from the following regression for columns 2 and 3 (with only the dependent variable changing):

$$\begin{aligned}(Charged\ 280-290g)_{it} = \alpha_0 + \beta_1(After2010)_{it} + \\ \beta_2(After2010 \times \%BlackOrHispanic)_{it} + \%BlackOrHispanic_{it} + \epsilon_{it}\end{aligned}$$

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

Table A16. Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors

	Atty. with 5+ Cases			Atty. with 15+ Cases		
Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010						
	Below 280g	280-290g	Above 290g	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)	(4)	(5)	(6)
Atty. Bunches at 280-290g Post-2010	-0.1271***	0.1438***	-0.0167	-0.0674	0.1394***	-0.0720
(15+ cases post-2010)	(0.0484)	(0.0356)	(0.0281)	(0.0837)	(0.0500)	(0.0639)
Constant	0.9024***	0.0305***	0.0671***	0.8651***	0.0209**	0.1140*
	(0.0247)	(0.0058)	(0.0247)	(0.0610)	(0.0081)	(0.0615)
Observations	1,583	1,583	1,583	692	692	692
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010						
	Below 280g	280-290g	Above 290g	Below 280g	280-290g	Above 290g
	(7)	(8)	(9)	(10)	(11)	(12)
Atty. Bunches at 50-60g Pre-2010	-0.0665***	0.0467***	0.0198	-0.0863***	0.0611***	0.0252
(15+ cases pre-2010)	(0.0245)	(0.0169)	(0.0151)	(0.0263)	(0.0167)	(0.0178)
Constant	0.9258***	0.0335***	0.0407***	0.9466***	0.0153	0.0382***
	(0.0168)	(0.0111)	(0.0115)	(0.0172)	(0.0096)	(0.0139)
Observations	1,278	1,278	1,278	956	956	956

Notes. Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(Charged\,X - Yg)_i = \alpha_0 + \beta_1 AttyBunchesAt280g_i + \epsilon_i$$

where $AttyBunchesAt280g$ is equal to one if the prosecutor is classified as a bunching prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). These regressions are restricted to post-2010 cases and to prosecutors with 5+ cases post-2010 in columns 1-3 and with 15+ cases post-2010 in columns 4-6. Note, to avoid a mechanical relationship in column (2), I use leave-out-means to classify bunching attorneys. Coefficients in panel B are estimated from the following regression for each range:

$$(Charged\,X - Yg)_i = \alpha_0 + \beta_1 AttyBunchesAt50g_i + \epsilon_i$$

where $AttyBunchesAt50g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 50g definition (i.e. the fraction of their cases that are from 50-60g is above the average fraction of 50-60g cases post-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 50-60g is at or below the average fraction of 50-60g cases post-2010). These regressions are restricted to post-2010 cases and to prosecutors with 5+ cases pre-2010 in columns 7-9 and with 15+ cases pre-2010 in columns 10-12.
*** p<0.01, ** p<0.05, * p<0.1

Table A17. Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors

Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)
Pct. of Cases Bunched at 280-290g	-0.5046***	0.5272***	-0.0227
(Leaving out current case in calculation)	(0.1162)	(0.0717)	(0.0976)
Constant	0.8892***	0.0400***	0.0707**
	(0.0343)	(0.0078)	(0.0346)
Observations	971	971	971
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Pct. of Cases Bunched at 50-60g	-0.3693***	0.2667***	0.1026*
(Leaving out current case in calculation)	(0.1042)	(0.0705)	(0.0565)
Constant	0.9196***	0.0356***	0.0448***
	(0.0123)	(0.0083)	(0.0087)
Observations	1,135	1,135	1,135

Notes. Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(Charged\,X-Yg)_i = \alpha_0 + \beta_1 PctBunching280g_i + \epsilon_i$$

where $PctBunchingAt280g$ is equal to the prosecutor's fraction of cases at 280-290g post-2010 (excluding the current observation) minus the average fraction of cases at 280-290g pre-2010. These regressions are restricted to post-2010 cases and to prosecutors with 10+ cases post-2010. Coefficients in panel B are estimated from the following regression for each range:

$$(Charged\,X-Yg)_i = \alpha_0 + \beta_1 PctBunching50g_i + \epsilon_i$$

where $PctBunchingAt50g$ is equal to the prosecutor's fraction of cases at 50-60g pre-2010 minus the average fraction of cases at 50-60g post-2010. These regressions are restricted to post-2010 cases and to prosecutors with 10+ cases pre-2010.
*** p<0.01, ** p<0.05, * p<0.1

Table A18. Persistence of Attorney-level Bunching Across Districts, from Analysis of Movers

	Pr(Atty. Bunches at 10-Year Mandatory Minimum in 2nd District)			
	(1)	(2)	(3)	(4)
Atty. Bunches at 10-Year MM in 1st District	0.184*	0.162**	0.263**	0.154*
	(0.0936)	(0.0816)	(0.108)	(0.0829)
Constant	0.500***	0.432***	0.462***	0.440***
	(0.0700)	(0.0580)	(0.0809)	(0.0577)
Bunching classification	280-290g, National	280-290g, Missing=0, National	280-290g, District	280-290g, Missing=0, District
Observations	109	148	79	144

Notes. Robust standard errors are in parentheses. The estimates in this table are based on the EOUSA data. For this analysis, I identify the attorneys who switch districts at some point in their career (using their initials recorded in the EOUSA case management system). I then identify the set of those attorneys who bunch at a 10-year mandatory minimum in their first district. I also limit the sample to attorneys who have at least 5+ cases in their first district and 5+ cases in their second district (this maintains the 10+ restriction but spreads it evenly across districts). Since I am analyzing movers, it is almost always the case that the cases in their first district are pre-2010 cases, meaning that the bunching classification is determined based on bunching at 50-60g. Finally, I regress an indicator equal to one if the attorney bunches at the 10-year threshold in their second district on whether they bunched at the 10-year threshold in their first district. I do this for four methods of classifying bunching attorneys. Columns 1 and 2 are detailed in Table A15. Columns 3 and 4 mirror those two approaches but define the "baseline" bunching at the district-level. For example, an attorney **i** in district **A** is defined as bunching at 50-60g in column 3 if their fraction of cases at 50-60g pre-2010 is above the fraction of cases at 50-60g in district **A** post-2010. In all cases, I find that an attorney who bunches above the mandatory minimum threshold in their first district is more likely to do so in their second district than an attorney who does not bunch above the mandatory minimum threshold in their first district.
*** p<0.01, ** p<0.05, * p<0.1

Table A19. Relationship between Various Bunching Ranges, Attorneys

	28-29g (1)	28-29g (2)	50-60g (3)	280-290g (4)	280-290g (5)	280-290g (6)
Atty. Bunches at 280-290g Post-2010	0.144**	0.140**	0.182***			
	(0.0625)	(0.0590)	(0.0664)			
Atty. Bunches at 28-29g Post-2010				0.155***	0.0876**	
				(0.0544)	(0.0340)	
Atty. Bunches at 50-60g Pre-2010						0.0575***
						(0.0172)
Constant	0.131***	0.120***	0.155***	0.0826***	0.0479***	0.0233**
	(0.0241)	(0.0232)	(0.0288)	(0.0271)	(0.0149)	(0.0105)
Sample Years	2011-2017	2011-2017	2000-2010	2011-2017	2011-2017	2011-2017
Sample Restriction	0-280g	0-280g, 290-1000g	0-1000g	29-1000g	0-28g, 29-1000g	0-1000g
Observations	843	910	1,976	483	840	1,135

Notes. Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Columns 1-3 estimate the likelihood an attorney who bunches at 280-290g (i.e. who has a fraction of cases at 280-290g post-2010 that is above the average fraction of 280-290g cases pre-2010) also bunches at 28-29g post-2010, 28-29g post-2010, and 50-60g pre-2010, respectively. Column 1 limits the sample to cases with below 280g to avoid a mechanical relationship. Column 2 does this by excluding only the 280-290g range from the sample. Both approaches yield similar results. Column 3, since the dependent variable is based on pre-2010 data, uses the full range of cases (0-1000g). Columns 4-6 estimate the likelihood an attorney who bunches at 28-29g post-2010 or 50-60g pre-2010 also bunches at 280-290g post-2010. As before, columns 4 and 5 exclude the 28-29g range to avoid a mechanical relationship. 28-29g is relevant post-2010 because 28g is the threshold for the 5-year mandatory minimum after 2010. 50-60g is relevant pre-2010 because 50g is the threshold for the 10-year mandatory minimum prior to 2010. All regressions in this table use the sample of attorneys who have 10+ cases (post-2010 for columns 1-5; pre-2010 for column 6). In all cases, an attorney who bunches at one mandatory minimum threshold is more likely to bunch at a separate mandatory minimum threshold.

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

Table A20. Bunching Attorneys and Case Complexity

	Neither District nor National Priority	Multiple Agencies	Any Gang Defendants	Retained Counsel	Multiple Opposing Counsel	Number of Court Events
	(1)	(2)	(3)	(4)	(5)	(6)
Atty. Bunches at 280-290g Post-2010	0.0656	-0.0063	0.0070	0.0587	-0.0595	0.0835
	(0.0938)	(0.0273)	(0.0267)	(0.0467)	(0.0562)	(0.0969)
Constant	0.1938***	0.0832***	0.0457***	0.1672***	0.3955***	2.025***
	(0.0428)	(0.0188)	(0.0120)	(0.0203)	(0.0347)	(0.0495)
Observations	2,330	2,548	2,548	818	910	2,587

Notes. Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Columns 1 estimates the likelihood an attorney who bunches at 280-290g (i.e. who has a fraction of cases at 280-290g post-2010 that is above the average fraction of 280-290g cases pre-2010) has a case that is classified as neither a district nor a national priority. Column 2-3 estimates whether those attorneys are more likely to have cases which involve multiple agencies or involve a defendant labeled as a gang defendant. Column 4-5 estimates whether those attorneys are more likely to have cases in which the defendant retains private counsel or has multiple opposing counsel working on the case. Note, however, that data on opposing counsel is missing for the vast majority of cases in this data. Column 6 estimates whether those attorneys have cases with more court events (e.g. new filings, indictments, information, superseding indictments, etc.).
*** p<0.01, ** p<0.05, * p<0.1

Table A21. Bunching at 280-290g and Number of Defendants or Witnesses

	Pr(280-290g)		
	(1)	(2)	(3)
After 2010	0.0532***	0.0552***	0.0773***
	(0.0054)	(0.0051)	(0.0057)
After 2010 x 2+ Defendants	0.1197***		
	(0.0176)		
After 2010 x 3+ Defendants		0.2289***	
		(0.0290)	
After 2010 x Any Witness			0.3933***
			(0.1212)
Constant	0.0023***	0.0030***	0.0039***
	(0.0004)	(0.0005)	(0.0005)
Observations	18,795	18,795	18,795

Notes. Robust standard errors are in parentheses. The estimates in this table are based on the EOUSA data. Column 1 shows that cases with 2 or more defendants are more likely to be recorded in the 280-290g range. Column 2 shows that cases with 3 or more defendants are also more likely to be recorded in the 280-290g range. Finally, column 3 shows that cases with witnesses are more likely to be recorded in the 280-290g range. Qualitative accounts (e.g. Lynch 2016) suggest that these decisions are discretionary and that they are ways prosecutors can increase the amount used at sentencing (e.g. charging defendants together as part of a conspiracy or using witness testimony about relevant conduct).
*** p<0.01, ** p<0.05, * p<0.1

Table A22. Bunching at 280-290g and Drug Conspiracy Charges

	Pr(Lead Charge = Conspiracy)		
	(1)	(2)	(3)
Case recorded at 280-290g	0.396***	0.307***	0.249***
	(0.0326)	(0.0329)	(0.0361)
Constant	0.166***	0.255***	0.314***
	(0.00279)	(0.00487)	(0.0156)
Sample restriction	0-1000g	50-1000g	280-1000g
Observations	18,062	8,236	1,116

Notes. Robust standard errors are in parentheses. The estimates in this table are based on the EOUSA data. The dependent variable is an indicator equal to one if the lead charge on the case is a drug conspiracy charge. Drug conspiracy charges are a tool that prosecutors can use to increase the weight involved in the offense because the total weight of the conspiracy is applied to each offender deemed involved in the conspiracy. The independent variable is whether the case involves 280-290g. Cases with 280-290g are substantially more likely to carry a lead conspiracy charge. This is true even when limiting to cases with 280-1000g only (see column 3).
*** p<0.01, ** p<0.05, * p<0.1

Table A23. Testing for Alleyne Effect in USSC Data, Based on Sentence Date

	Pr(Below 280g)			Pr(280-290g)			Pr(Above 290g)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sentenced After June	-0.0593**	-0.0454***	-0.0306**	0.0418***	0.0282**	0.0180	0.0175	0.0172	0.0126
	(0.0275)	(0.0157)	(0.0147)	(0.0152)	(0.0114)	(0.0115)	(0.0250)	(0.0140)	(0.0117)
Sentenced After June x Year = 2013	0.0532	0.0363	0.0190	-0.0519**	-0.0367**	-0.0256	-0.00131	0.000441	0.00663
	(0.0420)	(0.0216)	(0.0206)	(0.0240)	(0.0174)	(0.0173)	(0.0367)	(0.0211)	(0.0196)
Constant	0.852***	0.854***	0.844***	0.0344***	0.0372***	0.0431***	0.114***	0.109***	0.113***
	(0.0163)	(0.00879)	(0.0106)	(0.00807)	(0.00614)	(0.00796)	(0.0153)	(0.00925)	(0.00856)
Years Included	2011-2013	2011-2014	2011-2015	2011-2013	2011-2014	2011-2015	2011-2013	2011-2014	2011-2015
Observations	4,817	6,548	7,965	4,817	6,548	7,965	4,817	6,548	7,965

Notes. Robust standard errors in parentheses for columns 1, 4, and 7 (given the small number of total clusters). Standard errors clustered at the sentence year-month level in parentheses for columns 2-3, 5-6, and 8-9. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. In the EOUSA data, the drug quantity field does not necessarily display the final amount used for sentencing in the case. Instead, it is the amount entered into the case management system. Based on the user manual for the system, quantity is not a required field that must be updated throughout the life of the case. The user is first prompted to enter quantity when the case is opened. In practice, this is correlated with the final amount used in sentencing at the district-by-month level, and the patterns of bunching are similar in the EOUSA and USSC data. However, this difference is important for the analysis of *Alleyne*. In Table 7, I examine how bunching differs based on when the case is received since that is when prosecutors are likely recording the amount in the case management system. In cases received and recorded before *Alleyne*, prosecutors often indicate that the quantity involved is at 280g. In cases received and recorded shortly after *Alleyne*, prosecutors are less likely to indicate that 280g were involved. *Alleyne*, however, will affect the **final** amount used at sentencing in cases received both before and after the decision as long as they are sentenced after *Alleyne*. For cases initiated before *Alleyne* but sentenced after, prosecutors will likely need to file a superseding indictment or superseding information that addresses drug quantity. This means that there should also be a discontinuity in bunching in final amount used at sentencing for cases sentenced before and after *Alleyne*. The USSC data does not include exact date of sentencing, but it does include sentence month and year. I use this data in this table and show that bunching at final sentencing also falls, as we should expect, for cases sentenced shortly after *Alleyne*. Columns 1-3 examine the probability a case is recorded below 280g; columns 4-6, in 280-290g; and columns 7-9, above 290g. Columns 1, 4, and 7 estimate this using the year the case is decided and the two years before it; columns 2, 5, and 8 include the year of, the two years before, and one year after; columns 3, 6, and 9 include all years after 2010. I don't include cases sentenced in 2016 because the data does not include any cases sentenced after September 2016.

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

Table A24. Effect of *Alleyne v. US*, Accounting for Missing Values

	Pr(Case's Drug Weight is Missing)	Pr(Case is Charged with 280-290g, Missing = 0)
	(1)	(2)
After June 17th, 2011-2016	-0.0211	0.00438
	(0.0309)	(0.00869)
After June 17th, 2013	-0.0219	-0.0389*
	(0.0702)	(0.0223)
Constant	0.834***	0.0243
	(0.0690)	(0.0269)
Bandwidth	±150 days	±150 days
Observations	6,182	6,182

Notes. Standard errors clustered at the date the case is received in parentheses. The estimates in this table are based on the EOUSA data. The coefficients above are estimated from the following regression discontinuity style model:

$$Y_{it} = \alpha_0 + \beta_1 AfterJune17_{it} + \beta_2 DaysFrom_{it} + \beta_3 (AfterJune17 \times DaysFrom)_{it} + \delta_1 (AfterJune17 \times Year2013)_{it} + \delta_2 (DaysFrom \times Year2013)_{it} + \delta_3 (AfterJune17 \times DaysFrom \times Year2013)_{it} + D_{it} + \epsilon_{it}$$

where $AfterJune17$ is a dummy variable equal to one for cases received after June 17th in each year, $DaysFrom$, the running variable, is the date the case was received centered at zero on June 17th, and $Year2013$ is equal to one for cases received in 2013 (the year *Alleyne* is decided). In addition, all specifications above include day-of-week fixed effects, D_{it}, for the day the case is received. In column 1, Y_{it} is equal to one if the observation has a missing drug weight and equal to zero otherwise. There is little effect of *Alleyne* on the likelihood an observation has missing drug weight. In column 2, Y_{it} is equal to one if the drug weight is equal to 280-290g and equal to zero if the weight is outside the 280-290g range or if the weight is missing. There is still a decrease in bunching after *Alleyne* when accounting for missing values. Robustness to alternative ways of dealing with missing values is further explored in Appendix E.
*** p<0.01, ** p<0.05, * p<0.1

Table A25. Degree of Bunching Post-2010 by Race and District-level Characteristics

	Pr(280-290g)											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
After '10 x W	0.0187***	0.0222**	0.0136	0.0234**	0.0145	0.0198**	0.0155*	0.0084	0.0162*	0.0130**	0.0158**	0.0157**
	(0.0073)	(0.0105)	(0.0084)	(0.0107)	(0.0090)	(0.0085)	(0.0089)	(0.0062)	(0.0096)	(0.0062)	(0.0072)	(0.0067)
After '10 x BH	0.0397***	0.0439***	0.0308***	0.0505***	0.0305***	0.0353***	0.0365***	0.0360***	0.0359***	0.0301***	0.0353***	0.0364***
	(0.0028)	(0.0035)	(0.0028)	(0.0036)	(0.0027)	(0.0029)	(0.0033)	(0.0033)	(0.0031)	(0.0050)	(0.0088)	(0.0082)
After '10 x W x Char.	-0.0160	-0.0143	-0.0006	-0.0171	-0.0019	-0.0147	-0.0065	0.0086	-0.0069	0.0052	-0.0101	-0.0157**
	(0.0098)	(0.0122)	(0.0111)	(0.0120)	(0.0113)	(0.0106)	(0.0111)	(0.0116)	(0.0115)	(0.0255)	(0.0105)	(0.0067)
After '10 x BH x Char.	-0.0154***	-0.0171***	0.0082*	-0.0323***	0.0093**	-0.0021	-0.0023	-0.0012	-0.0011	0.0311	-0.0033	-0.0124
	(0.0043)	(0.0044)	(0.0044)	(0.0043)	(0.0044)	(0.0043)	(0.0045)	(0.0045)	(0.0045)	(0.0320)	(0.0113)	(0.0102)
Constant	0.0017	0.0015	0.0049***	0.0028**	0.0053***	0.0028**	0.0022*	0.0022**	0.0030**	0.0030**	0.0034**	0.0035***
	(0.0012)	(0.0011)	(0.0019)	(0.0014)	(0.0019)	(0.0012)	(0.0013)	(0.0011)	(0.0014)	(0.0012)	(0.0014)	(0.0013)
Characteristic	District-by-Year Above Med. # of Cases per Attorney	District Above Med. % of Guilty Cases	District Above Med. % of Declined Cases	District Above Med. % of Plea Cases	District Above Med. % of Cases Dismissed for 'Weak Evidence'	District Above Med. % of Cases Dismissed for 'Resources'	District Above Med. % Of Cases with Retained Counsel (based on '99-'02)	District Above Med. % Of Cases, Appointed Counsel	District Above Med. % Of Cases, Public Defender	District contains one of the ten biggest cities based on population	District is one of the ten biggest districts based on cases in 1999-2010	State has more than 4 gang members per 1,000 people (DEA 2009)
P-value: W = BH	0.0071	0.0505	0.0520	0.0158	0.0886	0.0852	0.0261	0.0001	0.0508	0.0414	0.0802	0.0425
P-value: W+Char. = BH+Char.	0.0035	0.0048	0.0012	0.0518	0.0004	0.0001	0.0006	0.0837	0.0003	0.0001	0.0000	0.0003
Observations	50,273	50,273	50,273	50,273	50,273	50,273	46,950	46,950	46,950	50,273	50,273	50,273

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data (although the EOUSA data is used to define characteristics of interest for columns 1-6). See Table 1 for notes about data construction. "Characteristic" or "Char." represents a dummy variable that is a district or district-by-year characteristic. The specific characteristic of interest is noted in the "Characteristic" row. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "P-value: W+Char. = BH+Char." reports the p-value from a test of the null hypothesis that the combined coefficients on "(After 2010 x White)+(After 2010 x White x Characteristic)" is equal to the combined coefficients on "(After 2010 x Black or Hispanic)+(After 2010 x Black or Hispanic x Characteristic)." Column 1 interacts the after 2010 by race dummy variables with a district-by-year dummy variable indicating if the district received above the median number of cases (per attorney) in the year. Column 2 studies districts above/below the median for percent of cases that end in a guilty verdict, column 3 studies districts above/below the median for percent of cases declined, and column 4 studies districts above/below the median for percent of cases that end in plea deals. Columns 5 and 6 study districts above/below the median for percent of cases declined due to "weak evidence" or "lack of resources" (as coded in the EOUSA case files, although codes are not present for all cases). Columns 7-9 use the USSC data from 1999-2002 on type of defense counsel to examine heterogeneity by type of defense counsel used in the district. Fourteen districts do not report this data in 1999-2002, hence the different observation count in these columns. Places with different rates of retained, appointed, or public defender defense counsel from 1999-2002 nevertheless have similar bunching at 280g post-2010. Column 10 examines heterogeneity based on whether the district contains one of the 10 largest cities in the US. Column 11 examines heterogeneity based on whether the district is one of the top 10 districts in terms of number of crack-cocaine cases from 1999-2010. Column 12 uses the DEA Drug Threat Assessment to classify states based on gang activity and examine heterogeneity by whether the state has more or less than 4 gang members per 1,000 people. I explore heterogeneity by additional district-level characteristics in Tables A29-30.
*** p<0.01, ** p<0.05, * p<0.1

Table A26a. Relationship between Race and Gang Involvement among Federal Inmates

	In Drug Org. Prior to Arrest	Illegal Income Prior to Arrest	In Drug Org.	Hired Private Counsel
	(1)	(2)	(3)	(4)
Panel A. Inmates Reporting Any Drug Offense				
Black or Hispanic	-0.121***	-0.114***	-0.134***	-9.99e-06
	(0.0251)	(0.0326)	(0.0407)	(0.0311)
Constant	0.222***	0.513***	0.292***	0.363***
	(0.0231)	(0.0282)	(0.0359)	(0.0268)
Sample Restriction	-	-	Illegal Income > 0	-
Observations	1,268	1,219	521	1,255
Panel B. Inmates Reporting a Crack-Cocaine Offense				
Black or Hispanic	-0.116	-0.102	-0.138	0.104
	(0.105)	(0.138)	(0.156)	(0.113)
Constant	0.200*	0.615***	0.250	0.214*
	(0.104)	(0.135)	(0.154)	(0.110)
Sample Restriction			Illegal Income > 0	
Observations	324	311	160	322

Notes. Robust standard errors in parentheses. The estimates in this table are based on the SIFCF data. Column 1 estimates the racial difference in whether an inmate reports being involved in a drug organization prior to arrest. Column 2 estimates the racial difference in whether an inmate reports earning illegal income prior to arrest. Column 3 estimates the racial difference in whether an inmate reports being involved in a drug organization prior to arrest, restricting to inmates that report positive illegal income prior to their arrest. Column 4 estimates the racial difference an inmate reports retaining private counsel. Panel A restricts the analysis to all drug offenders, and panel B restricts to offenders that report involvement with crack-cocaine.
*** p<0.01, ** p<0.05, * p<0.1

Table A26b. Relationship between Race and Drug Involvement among Federal Inmates

	(1) Weight	(2) Weight	(3) Pr(Above 280g)	(4) Pr(Above 280g)	(5) # of Members in Gang	(6) Leader in Gang	(7) Middle Man	(8) Underling	(9) Seller	(10) Other Role
Black or Hispanic	18.5	-65.8	0.00883	0.0221	0.936	-0.0136	-0.118*	0.110*	0.0683	-0.0464
	(32.2)	(701.7)	(0.0825)	(0.113)	(4.53)	(0.0616)	(0.0653)	(0.0641)	(0.0669)	(0.0533)
Constant	91.1***	862.3	0.0833	0.214*	18.0***	0.194***	0.278***	0.167***	0.208***	0.153***
	(30.0)	(681.6)	(0.0801)	(0.110)	(3.51)	(0.0469)	(0.0531)	(0.0442)	(0.0482)	(0.0427)
Sample Restriction	0-1000g	0-25000g	0-1000g	0-25000g	In Drug Org. = 1	In Drug Org. = 1	In Drug Org. = 1	In Drug Org. = 1	In Drug Org. = 1	In Drug Org. = 1
Observations	229	272	229	272	154	166	166	166	166	166

Notes. Robust standard errors in parentheses. The estimates in this table are based on the SIFCF data. This table summarizes drug involvement by race for inmates reporting crack-cocaine involvement. Column 1-4 estimate racial differences in reported weight involved or whether that weight was above 280g. These regressions are restricted to observations with non-missing weight. Columns 5-10 estimate racial differences in the inmates' reported role in the gang, restricting to inmates who report involvement in a drug organization prior to arrest. Column 5 estimates racial differences in the size of the gang and columns 6-10 estimate racial differences in whether the inmate reports being a leader, middle man, underling, seller, or other role.
*** p<0.01, ** p<0.05, * p<0.1

Table A27. Relationship between Bunching at 280g and Judge Characteristics

	Pr(280-290g)	Pr(280-290g)	Pr(280-290g)
	(1)	(2)	(3)
After 2010	0.0898***	0.0852***	0.0974***
	(0.0119)	(0.0240)	(0.0222)
After 2010 ×White Judge		-0.0018	
		(0.0023)	
After 2010 ×Republican Judge			-0.0013
			(0.0015)
Constant	0.0038***	0.0053**	0.0046***
	(0.0007)	(0.0022)	(0.0013)
Observations	8,794	8,794	8,794

Notes. Standard errors clustered at the judge level in parentheses. The estimates in this table are based on the EOUSA data. This table uses data on the first judge listed in the case file. Approximately 70% of crack-cocaine cases in the EOUSA data have one judge listed; 25% have two listed; and 5% have more than two listed. Results are robust to using the last judge listed on the case (available upon request). I can match judge race and political party to approximately half of the cases in the EOUSA data. For data on judge characteristics, I use the file provided by Cohen and Yang (2019). I estimate whether bunching at 280g is related to judge race or judge political party. Column (1) shows that the level of bunching is similar for cases where I can match judge characteristics. Column (2) shows that judge race does not affect bunching at 280g. Column (3) shows that judge political party does not affect bunching at 280g.
*** p<0.01, ** p<0.05, * p<0.1

Table A28. Relationship between Various Bunching Ranges, Judges

	28-29g	28-29g	50-60g	280-290g	280-290g	280-290g
	(1)	(2)	(3)	(4)	(5)	(6)
Judge Bunches at 280-290g Post-2010	-0.00844	-0.00438	0.0536			
	(0.0303)	(0.0284)	(0.0415)			
Judge Bunches at 28-29g Post-2010				-0.00149	-0.0135	
				(0.0522)	(0.0328)	
Judge Bunches at 50-60g Pre-2010						0.0239
						(0.0213)
Constant	0.150***	0.138***	0.201***	0.168***	0.107***	0.0676***
	(0.0191)	(0.0181)	(0.0249)	(0.0388)	(0.0248)	(0.0178)
Sample Restriction	0-280g	0-280g, 290-1000g	0-1000g	29-1000g	0-28g, 29-1000g	0-1000g
Observations	767	825	2,686	468	790	1,261

Notes. Standard errors clustered at the judge level in parentheses. The estimates in this table are based on the EOUSA data. This table uses data on the first judge listed in the case file. Approximately 70% of crack-cocaine cases in the EOUSA data have one judge listed; 25% have two listed; and 5% have more than two listed. Results are robust to using the last judge listed on the case (available upon request). See Table A22 for a discussion of the dependent and independent variables in column 1-6. The major difference is that these regressions examine judges classified as "bunching" at a given range. This is possible because the EOUSA files contain a judge ID for many cases. I use that judge ID to calculate the fraction of cases at 280-290g post-2010, 28-29g post-2010, and 50-60g pre-2010 for each judge. 28-29g is relevant post-2010 because 28g is the threshold for the 5-year mandatory minimum after 2010. 50-60g is relevant pre-2010 because 50g is the threshold for the 10-year mandatory minimum prior to 2010. All regressions in this table use the sample of judges who have 10+ cases (post-2010 for columns 1-5; pre-2010 for column 6). Judges who bunch at one mandatory minimum threshold are not more likely to bunch at other mandatory minimum thresholds.

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

Table A29. Degree of Bunching Post-2010 by Race and Share Black in District Population

	Coefficient on After 2010 x Black or Hispanic	Number of Observations	Mean Share Black in District in 2010
Decile = 1	-0.00574	1,376	0.015
	(0.00945)		
Decile = 2	0.0420**	1,444	0.034
	(0.0179)		
Decile = 3	-0.00238	3,819	0.054
	(0.0163)		
Decile = 4	0.0241***	4,946	0.066
	(0.00626)		
Decile = 5	0.0156***	5,623	0.089
	(0.00542)		
Decile = 6	0.0116	4,961	0.121
	(0.0214)		
Decile = 7	-0.0245	7,297	0.151
	(0.0291)		
Decile = 8	0.0450**	7,284	0.207
	(0.0200)		
Decile = 9	0.0278	7,924	0.294
	(0.0292)		
Decile = 10	0.0403***	5,416	0.390
	(0.00635)		

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. In this table, I split districts into deciles based on the share of black people in the district in 2010 (data aggregated from county-level data from Opportunity Insights). These deciles are formed at the district level, which is why the case counts can differ dramatically across deciles. Each decile row shows estimates from separate regressions which all take the form:

$$\begin{aligned}(\text{Charged } 280-290g)_{it} = \alpha_0 + \beta_1(\text{After2010})_{it} + \\ \beta_2(\text{After2010}\times\text{BlackOrHispanic})_{it} + \\ \text{BlackOrHispanic}_{it} + \epsilon_{it}\end{aligned}$$

where β_2 identifies the racial disparity in bunching at 280-290g. In each regression, the sample is limited to only those districts in the decile bin. Column 1 displays the estimates, column 2 displays the observation counts, and column 3 displays the average share of the population that is black for each decile bin. There is no clear relationship between decile and the racial disparity in bunching. This suggests that prosecutors are not simply responding to potential juror bias; if so, we might expect bunching to decrease as the probability of the jury containing a black juror increases. Note that electorate preferences are unlikely to affect decisions in this context; federal judges are appointed and AUSAs are hired directly by US Attorney's offices.
*** p<0.01, ** p<0.05, * p<0.1

Table A30. Degree of Bunching Post-2010 by Race and Additional District-level Characteristics

	Pr(280-290g)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After 2010 x White	0.0095*	0.0100	0.0120	0.0156	0.0118*	0.0179	0.0129**
	(0.0054)	(0.0091)	(0.0101)	(0.0117)	(0.0070)	(0.0177)	(0.0059)
After 2010 x Black or Hispanic	0.0387***	0.0442***	0.0463***	0.0446***	0.0311***	0.0526***	0.0309***
	(0.0027)	(0.0037)	(0.0046)	(0.0044)	(0.0030)	(0.0065)	(0.0022)
Constant	0.0025**	0.0054**	0.0022	0.0011	0.0022**	-0.0000	0.0035***
	(0.0010)	(0.0022)	(0.0016)	(0.0011)	(0.0011)	(0.0000)	(0.0011)
Pre-2010 Difference	Pr(0-280g,W) > Pr(0-280g, BH)	Pr(50-280g,W) > Pr(50-280g, BH)	Pr(60-280g,W) > Pr(60-280g, BH)	(Mean Wgt., W) > (Mean Wgt., BH)	(Mean Wgt.>60g, W) > (Mean Wgt.>60g, BH)	Pr(Trial, W) > Pr(Trial, BH)	Pr(Trial, W) < Pr(Trial, BH)
P-value: W = BH	0.0000	0.0005	0.0019	0.0205	0.0115	0.0653	0.0040
Observations	35,537	19,262	15,062	15,777	24,945	6,969	43,344

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. Each column in this table limits to a subset of districts based on a different pre-2010 characteristic. Column 1 limits to districts where white offenders are relatively more likely to be charged with amounts below 280g prior to 2010. Column 2 limits to districts where white offenders are relatively more likely to be charged with amounts from 50-280g prior to 2010. Column 3 limits to districts where white offenders are relatively more likely to be charged with amounts from 60-280g prior to 2010. If we consider that cases bunched at 280-290g after 2010 are cases that would have been recorded below 280g prior to 2010, then these are districts in which white offenders have a larger base of offenders from which to bunch at 280-290g after 2010. Columns 4 presents a similar exercise by limiting to districts in which the mean drug weight for white offenders is higher than the mean drug weight for black and Hispanic offenders pre-2010. Column 5 makes a similar restriction, but only uses the mean of the weights above 60g. Finally, column 6-7 consider how differences in plea rates might affect racial disparities in bunching at 280-290g. Column 6 limits to districts where white offenders are relatively more likely to go to trial. Column 7 limits to districts where black and Hispanic offenders are relatively more likely to go to trial.
*** p<0.01, ** p<0.05, * p<0.1

Table A31. Relationship between Bunching in EOUSA and State-level Racial Animus

	280-290g	280-290g, Missing = 0	# of Attys in State who Bunch at 280g
	(1)	(2)	(3)
After 2010	0.0756***	0.0163***	-
	(0.0123)	(0.00287)	-
Above Med. Racial Animus	-0.00187	-0.000390	1.737**
	(0.00122)	(0.000447)	(0.690)
After '10 × Above Med. Racial Animus	0.00150	0.0106***	-
	(0.0138)	(0.00365)	-
Constant	0.00520***	0.00182***	-
	(0.00111)	(0.000388)	-
Observations	19,241	49,051	51

Notes. Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. See Table A15 for a discussion of the "280-290, Missing=0" dependent variable. Columns 1 and 2 interact the after 2010 dummy variable with a dummy variable equal to one when the state where the case is received is above the median level of racial animus and equal to zero if it is below the median level. Coefficients are estimated from the following regression for columns 1 and 2 (with only the dependent variable changing):

$$(Charged\ 280-290g)_{it} = \alpha_0 + \beta_1(After2010)_{it} + \beta_2(After2010 \times AboveMedRA)_{it} + AboveMedRA_{it} + \epsilon_{it}$$

Since racial animus is a measure that varies across districts, column 2 results are particularly noteworthy (using the "missing included" version of 280-290g accounts for some of the cross-district variation in drug weight reporting). Finally, column 3 estimates a state-level regression of the number of bunching attorneys in the state (defined as an attorney whose fraction of cases at 280-290g post-2010 is above the average fraction at 280-290g pre-2010) on the above median racial animus dummy variable.
*** p<0.01, ** p<0.05, * p<0.1

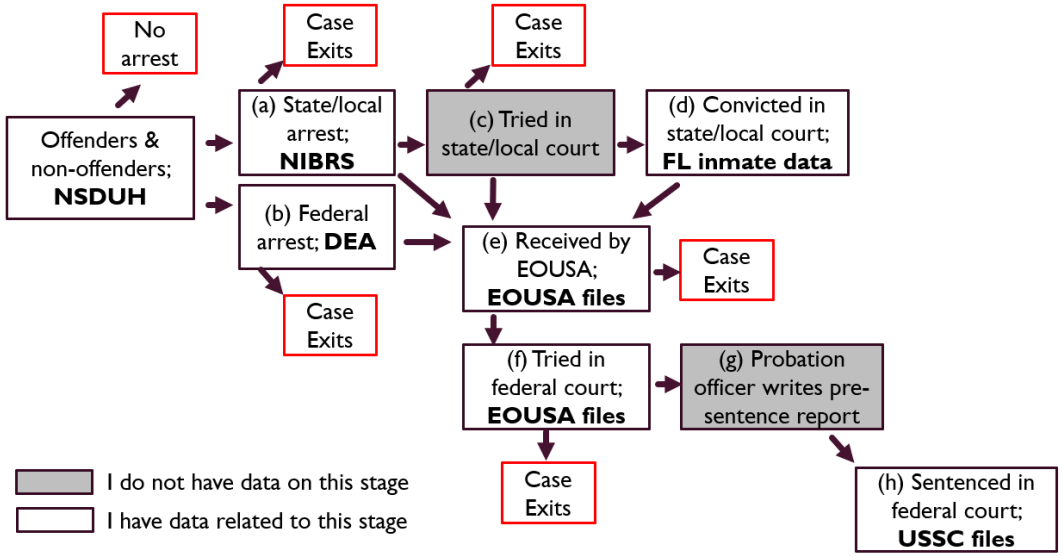
Table A32. Robustness Tests for Relationship between IAT and Racial Disparity in Bunching at 280g

	Pr(280-290g)	
	(1)	(2)
After 2010 x Above Med. IAT	-0.0293	-
	(0.0323)	-
After 2010 x Above Med. IAT x White	-	-0.0194
	-	(0.0156)
After 2010 x Above Med. IAT x Black or Hispanic	0.0332	0.0127
	(0.0294)	(0.00897)
Constant	0.0107***	-0.0443
	(0.00132)	(0.0328)
Observations	49,609	49,609

Notes. Standard errors clustered at the district level are in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. Both columns use a district level measure of animus, the implicit association test scores for lawyers (and other legal-service workers) aggregated to the district level. In this table, I use that score to classify districts as above median on the IAT (i.e. a greater implicit bias toward black people) or below median on the IAT. Column 1 follows Column 6 of Table 9 and includes state fixed effects interacted with the after 2010 by race dummy variables. Column 2 follows Column 4 of Table 9 and omits those fixed effects but includes individual controls (college, male, age, criminal history, citizenship, dependents, state caseload, the average black-white gap in years sentenced at the district level pre-2010, and the average black-white gap in drug weights charged above 50g at the district level pre-2010) interacted with the after 2010 by race dummy variables and district controls for economic characteristics (median household income in 2016, non-white share of population in 2010, population density in 2010, fraction with college in 2010, poor share in 2010, log of wage growth for high school graduates, black-white and Hispanic-white differences in incarceration and income conditional on parent income rank at the 25th percentile, job density in 2013, and annual job growth from 2004-2013) interacted with the after 2010 by race dummy variables.
*** p<0.01, ** p<0.05, * p<0.1

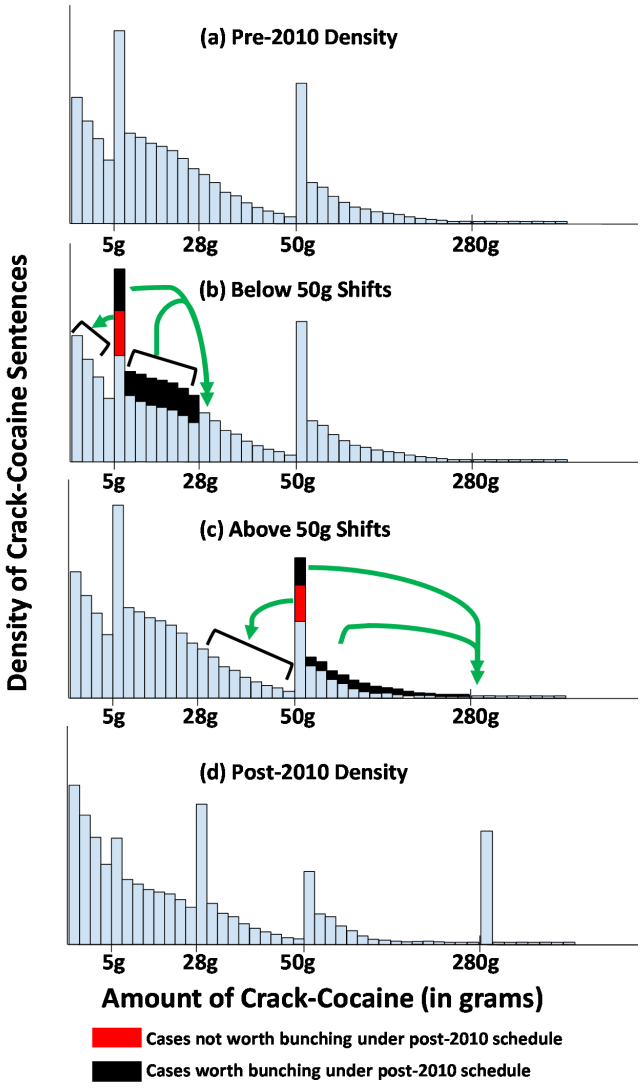
II. Figures

Figure A1. Graphical Illustration of Timeline from Arrest to Sentencing.



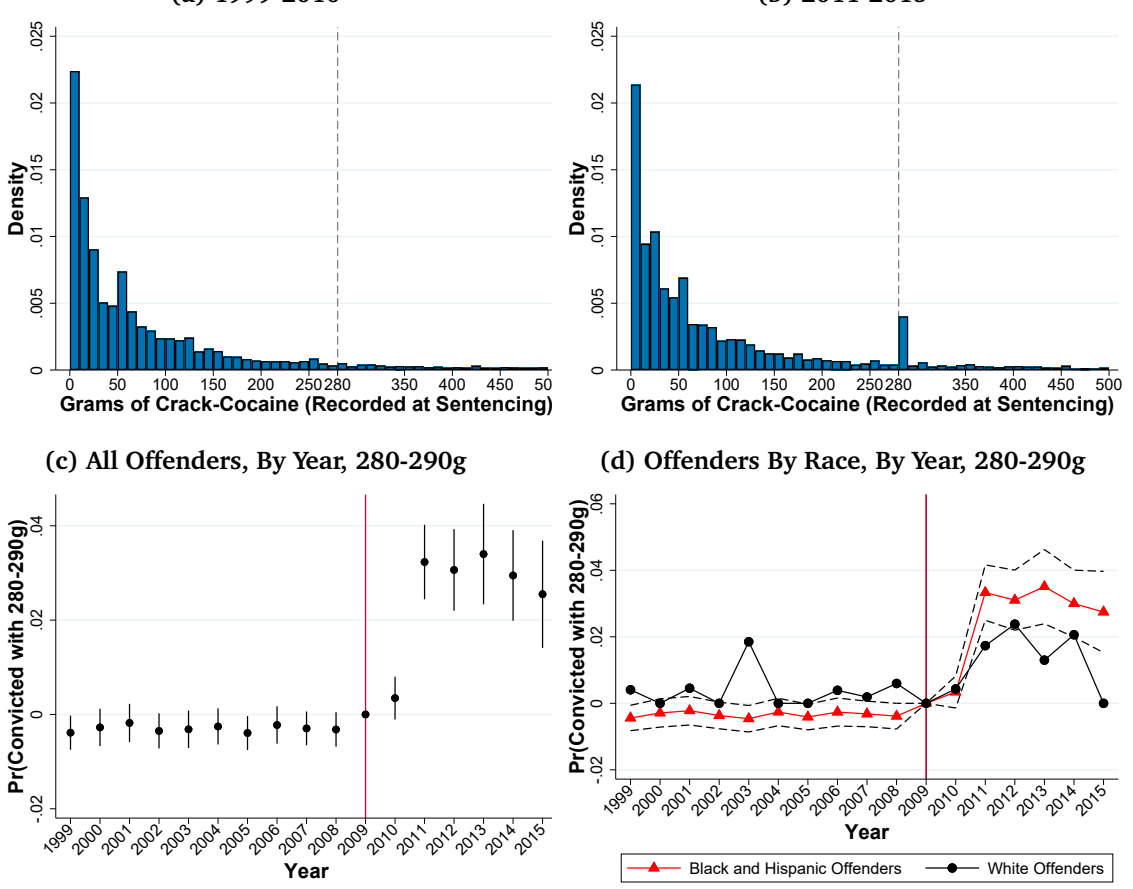
Notes. The figure above details the timeline from arrest to sentencing. Before arrest, the eventual arrestees come from the set of all people, some of whom are innocent and some of whom are guilty. Some individuals from this group are arrested by state/local police or federal agents. Of those arrested by state/local police, their case can be dismissed, prosecuted in state/local court, or passed on to federal authorities. Cases prosecuted in state/local court can leave the system if they are found not guilty, dismissed, etc., they can be convicted, or they can be sent to federal authorities. In fact, even cases convicted in state courts can be sent to federal authorities. Individuals arrested by federal agents are typically referred to the EOUSA directly. Once a case is received by the EOUSA, it can leave the system via a dismissal, declination, etc., or it can be prosecuted in federal court. For cases convicted in federal court, a probation officer prepares a pre-sentence report, and ultimately, the offender is sentenced. I have obtained data at nearly all of these steps. The two steps for which I lack data are in the middle of steps where bunching does not change, which suggests that nothing changes in the middle step. Note, these data sets are not linked, but observing the distributions in each separately is informative about where bunching at 280-290g first occurs.

Figure A2. Graphical Illustration of Conceptual Model, Prosecutor Responses to the FSA.



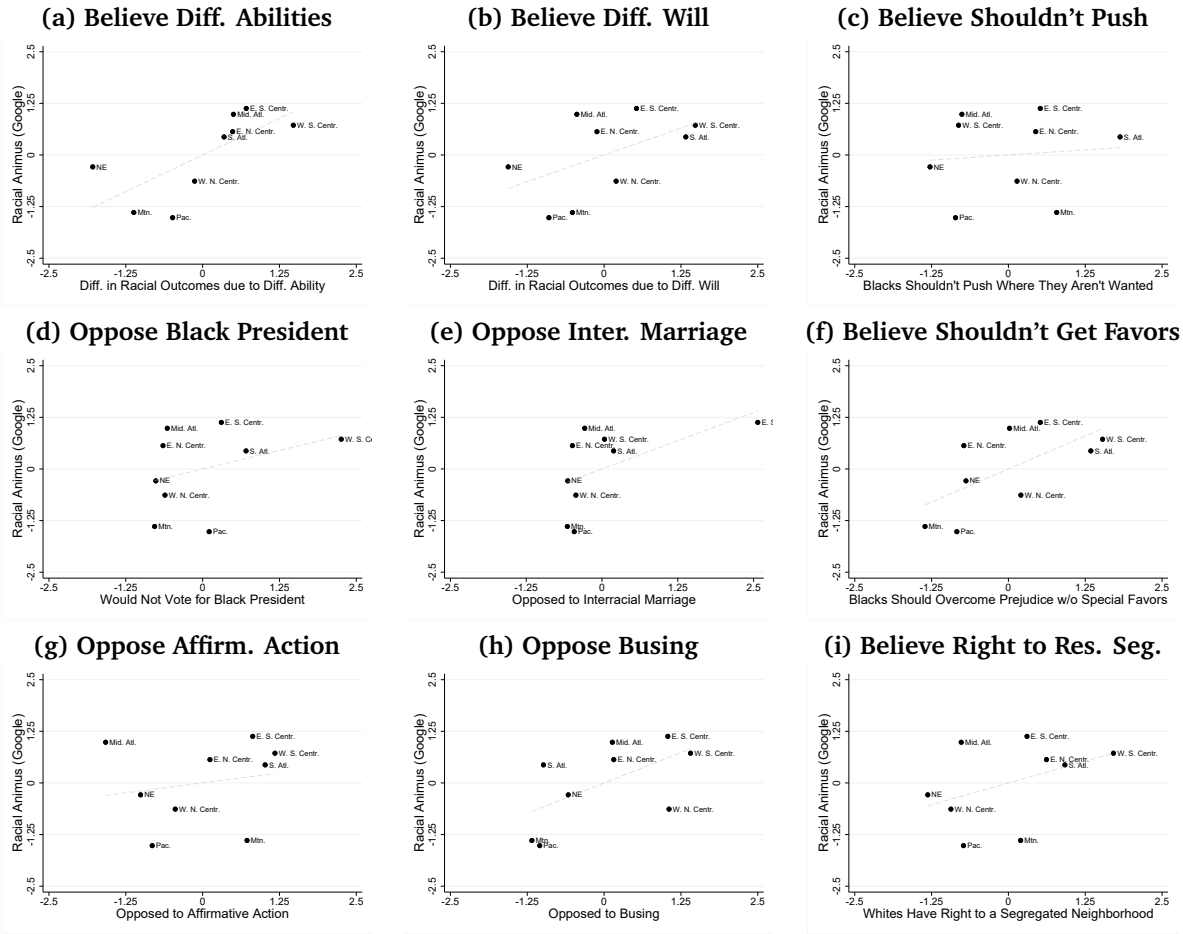
Notes. Panel (a) displays a hypothetical pre-2010 distribution of weights, with bunching at 5g and 50g due to round-number bias and prosecutor discretion. Panel (b) shows how the 0-5g, 5-28g, and 28-50g ranges will change after 2010. Some cases bunched at 5g will not be worth bunching at 28g (depicted in red), and they will shift into the 0-5g range. Some cases bunched at 5g and some cases from 5-28g will be worth bunching at 28g (depicted in black), and they will shift into the 28-50g range. Panel (c) illustrates a similar phenomena for the 50-280g range–some cases will shift down into the 28-50g range and some will shift up to the 280-290g range. Panel (d) shows the hypothetical post-2010 distribution of weights, with bunching at 5g and 50g due to round-number bias and bunching at 28g and 280g due to prosecutor discretion.

Figure A3. Changing Distribution of Drug Amounts Around 280g Pre- and Post-2010, USSC



Notes. Panels (a) and (b) plot the distribution of drug amounts recorded in federal crack-cocaine sentences starting at 0 grams and ending at 500 grams for 1999-2010 (when the mandatory minimum threshold was 50g) and 2011-2015 (when it was 280g). In panel (c), I estimate the main bunching coefficient by year (relative to 2010) and plot the coefficients with 90% confidence intervals. Panel (d) plots the coefficients and confidence interval for black and Hispanic offenders and the coefficients for white offenders (I do not include confidence intervals for white offenders because their estimates by year are extremely noisy). These plots are created from the USSC data. See Table 1 for notes about data construction.

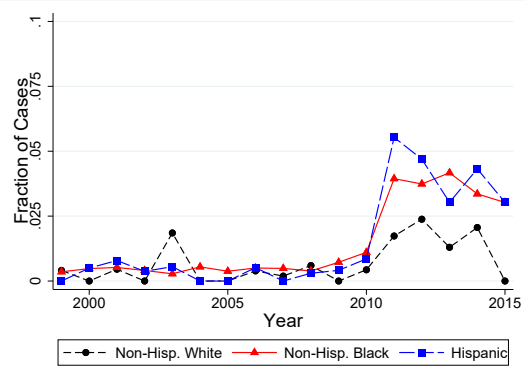
Figure A4. Relationship between Google Trends Racial Animus Measure and GSS Responses from Respondents with High Formal Education on Attitudes about Race



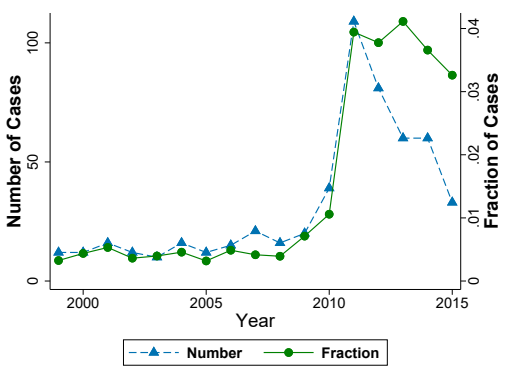
Notes. The figures above plot the relationship between the Google Trends racial animus measure (standardized and centered at zero) and various measures of attitudes about race from the General Social Survey (GSS) from 1972-2018 (not all questions are present in all years; also standardized and centered at zero). For the GSS measures, I limit the sample to respondents with a graduate degree or higher to test if the Google Trends racial animus measure is correlated with racial attitudes of people with high formal education. The public sample of the GSS only includes region identifiers. I aggregate the Google Trends measure to the region level by taking the mean across all states in the region. The regions are: Northeast, West North Central, Pacific, Mountain, East North Central, Mid Atlantic, South Atlantic, West South Central, and East South Central. The GSS questions are: Do you believe... (a) racial differences in outcomes are due to different abilities by race (available 1977-2018), (b) racial differences in outcomes are due to different will by race (1977-2018), (c) black shouldn't push where they aren't wanted (1972-2002), (f) blacks should overcome prejudice without special favors (1994-2018), and (i) whites have a right to a segregated neighborhood (1972-1996)? And are you opposed to... (a) voting for a black president (1972-2010), (b) interracial marriage (1972-2002), (c) affirmative action (1994-2018), (d) desegregation busing (1972-1996)?

Figure A5. Alternative Ways of Visualizing Bunching at 280-290g

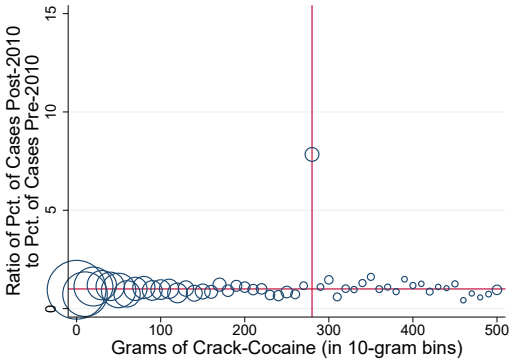
(a) Over Time by Race and Ethnicity



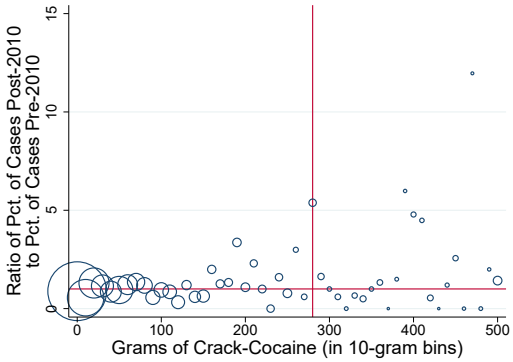
(b) Number and Share Over Time



(c) Bunching Ratio, 0-500g, Black and Hispanic



(d) Bunching Ratio, 0-500g, White

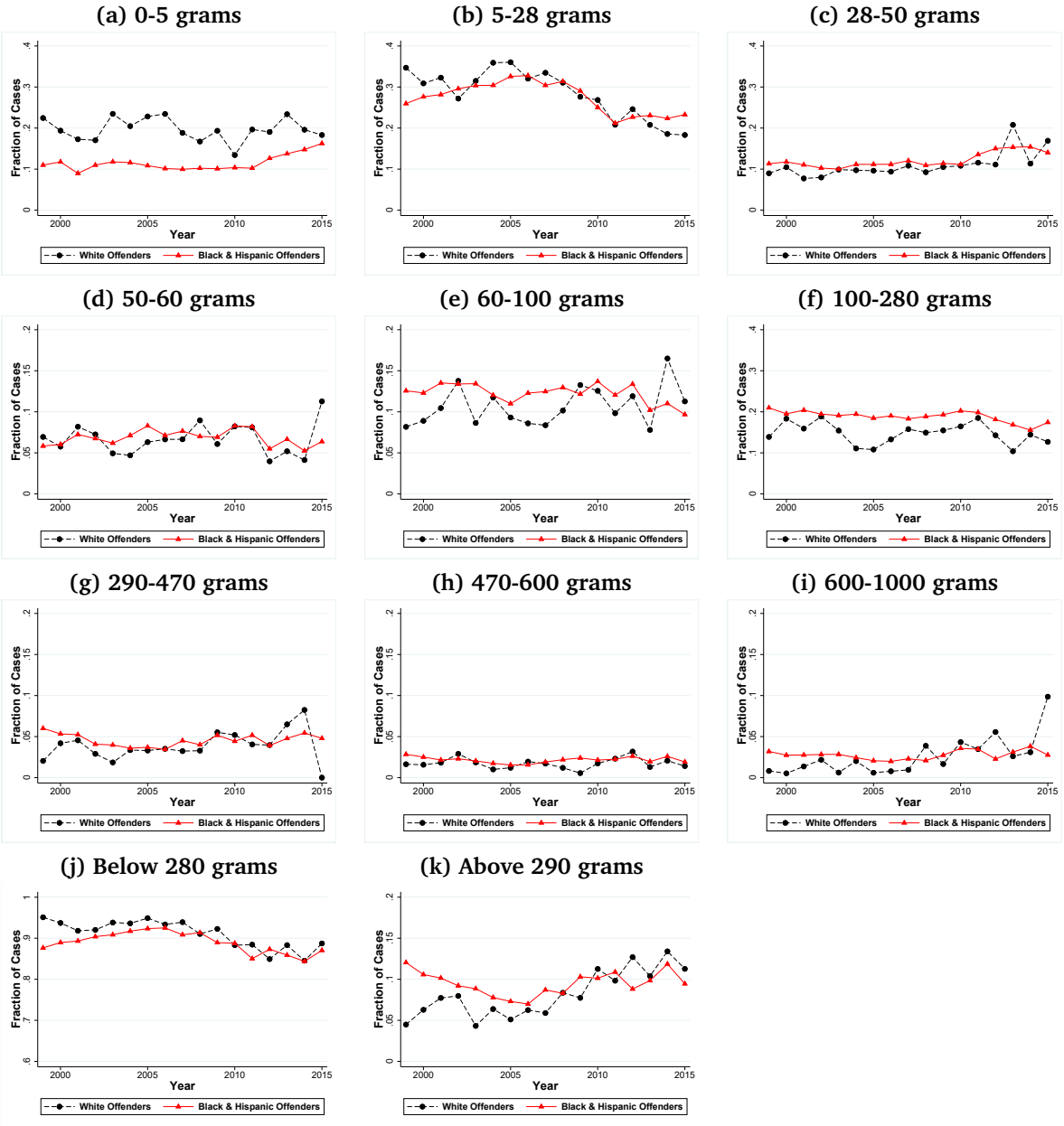


Notes. Panel (a) plots the fraction of cases bunched at 280-290g for non-Hispanic white offenders, non-Hispanic black offenders and Hispanic offenders. In addition, the sample used in this figure excludes any districts where greater than 20% of cases are missing ethnicity information. In this sample, pr(280-290g) increases by about 4.1 percentage points for Hispanic offenders and about 3 percentage points for non-Hispanic black offenders (p-value=0.14). The Southern District of NY has a large number of cases and many that are missing ethnicity information. Simply excluding that one district implies a 3.8 percentage point increase in pr(280-290g) for Hispanic offenders and a 3 percentage point increase for non-Hispanic black offenders (p-value=0.20). Panel (b) plots the total number of offenses with 280-290g over time and the share (or fraction) of cases with 280-290g over time. This plot is created from the USSC data. Panels (c) and (d) above plots the bunching ratio for each 10-gram bin from 0-500 grams by race. The bunching ratio for each bin b is defined as follows:

$$\text{Bunching Ratio}_b = \frac{\text{\% of cases in b post-2010}}{\text{\% of cases in b pre-2010}}$$

If the distributions are the same pre- and post-2010, the bunching ratio will equal 1 (marked by the horizontal red line). If the ratio is above 1, there is a higher degree of bunching in bin b post-2010. If the ratio is below 1, there is a lower degree of bunching post-2010. The size of the marker for each bin b is weighted by the total number of cases in the bin pre- and post-2010 (relative to rest of the group included in the plot, not relative to the full sample). These plots are created from the USSC data. See Table 1 for notes about data construction.

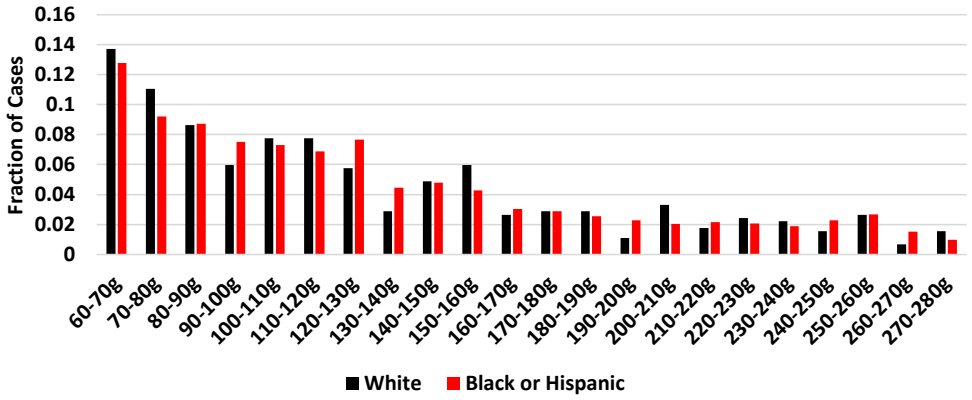
Figure A6. Changing Distribution of Drug Weights Over Time, By Race, USSC



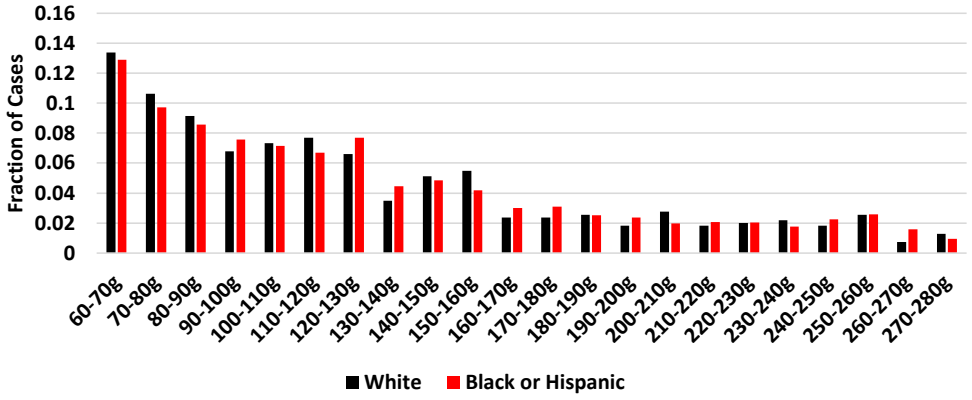
Notes. The figures above plot the share of cases in the specified range by year for white offenders and black and Hispanic offenders. For example, panel (a) plots the share of cases with 0-5g (not including 5g) in each year from 1999-2015. Panel (b) plots the share of cases with 5-28g in each year from 1999-2015, and so on. These plots are created from the USSC data. See Table 1 for notes about data construction.

Figure A7. Distribution of Pre-2010 Charged Amount by Race, 60-280g, by Characteristics

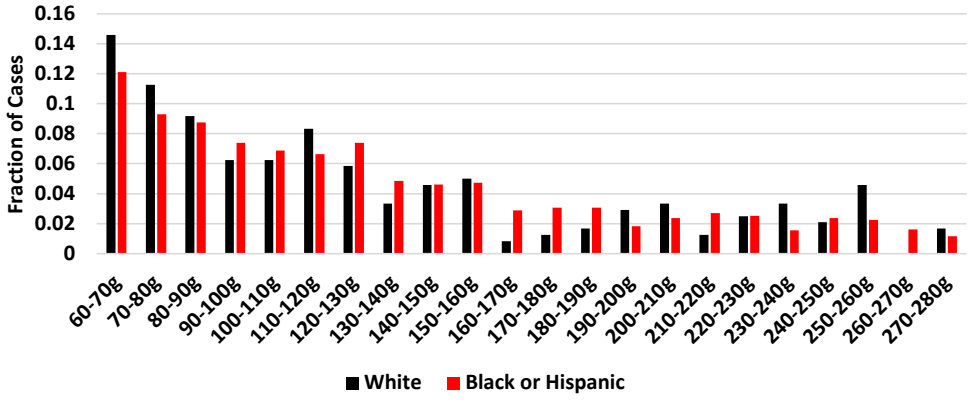
(a) Below Med. Criminal History



(b) Below Med. Predicted Sentence



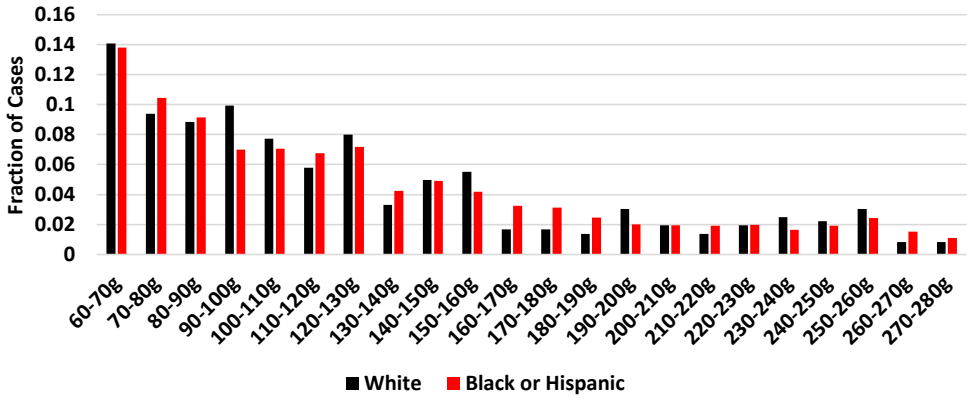
(c) Below Med. Racial Animus State



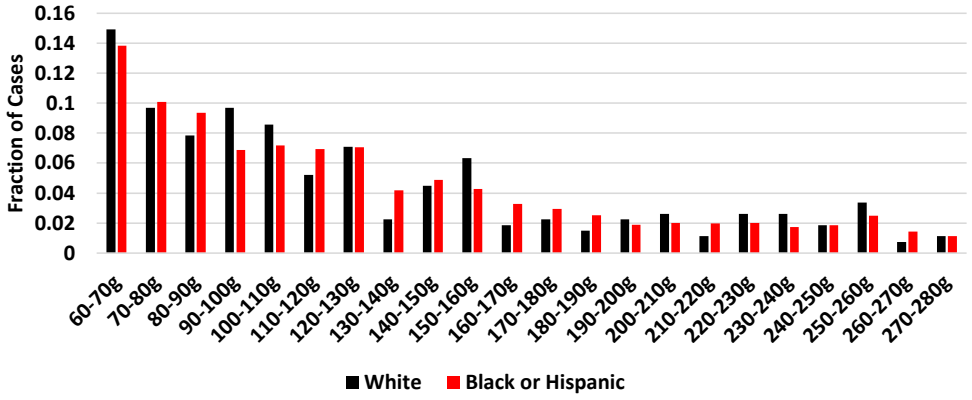
Notes. The figures above plot the distribution of charged amounts pre-2010 from 60-280g by race for various subsets of offenders. Panel (a) plots the distributions for offenders with below median criminal history scores. Panel (b) plots them for offenders with below median predicted sentenced (using exogenous factors, like age, sex, education, etc.). Panel (c) plots them for offenders convicted in states with below median levels of racial animus. In each plot, the distributions are similar. Kolomogorov-Smirnov tests of the equality of distributions fail to reject the null in all three cases: (a) p-value=0.60; (b) p-value=0.78; (c) p-value=0.45. This shows that the distributions of drugs charged from 60-280g are similar by race even within observably similar groups. These plots are created from the USSC data. See Table 1 for notes about data construction.

Figure A7. Distribution of Pre-2010 Charged Amount by Race, 60-280g, by Characteristics

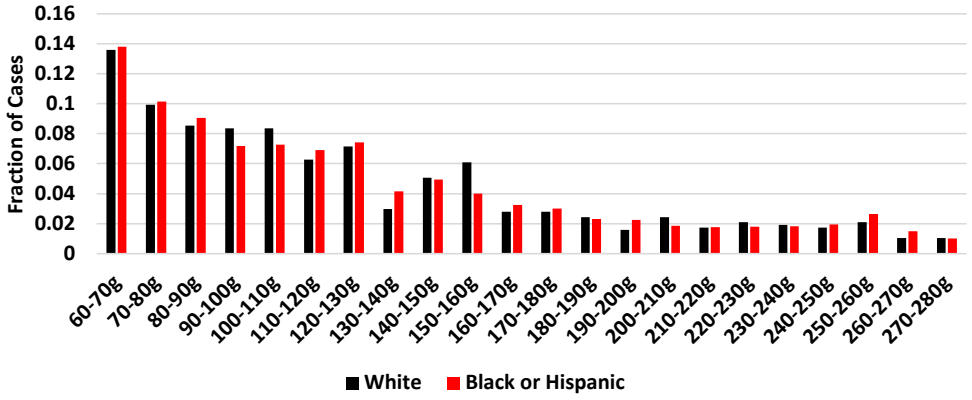
(d) Above Med. Criminal History



(e) Above Med. Predicted Sentence

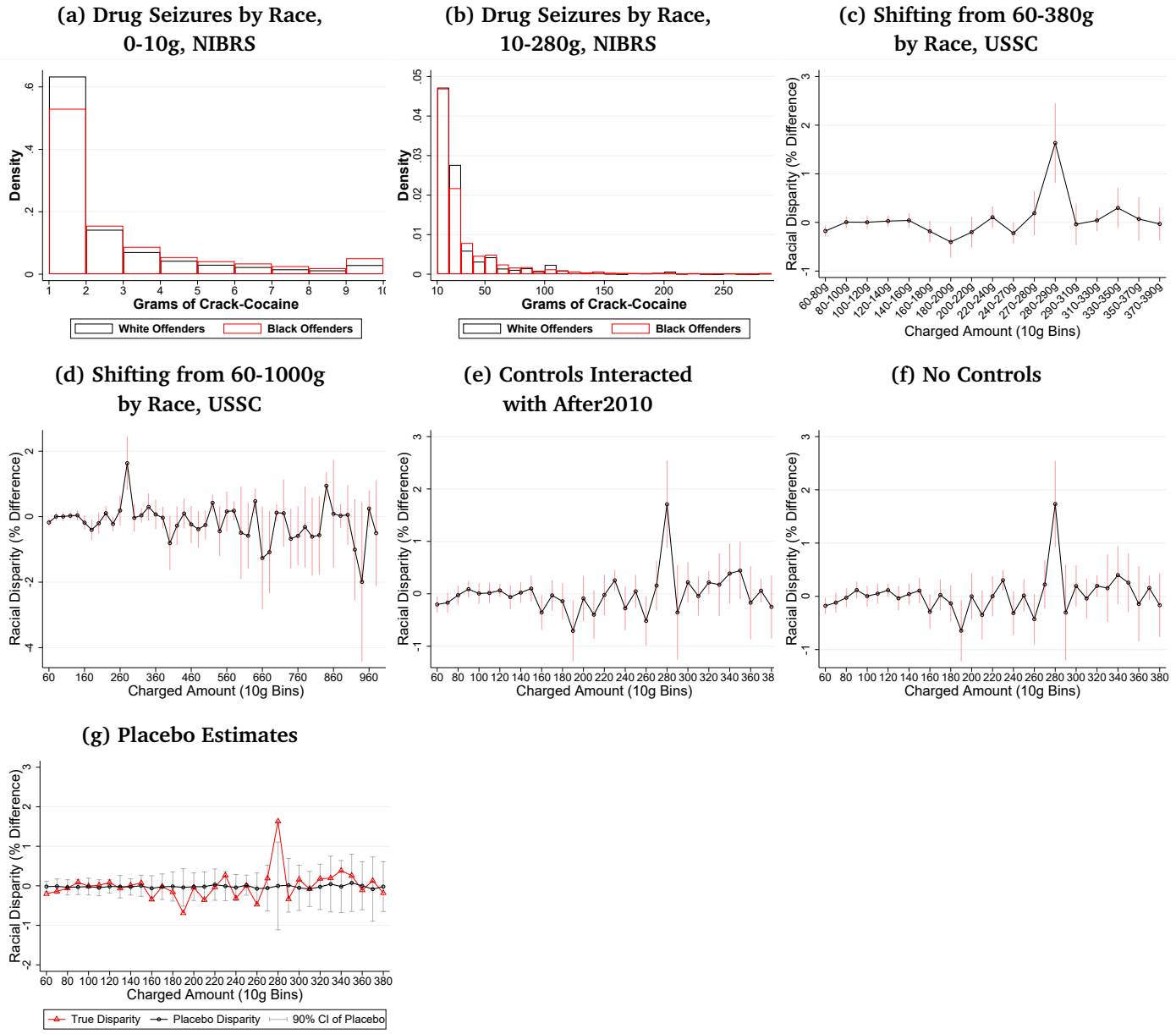


(f) Below Med. Racial Animus State



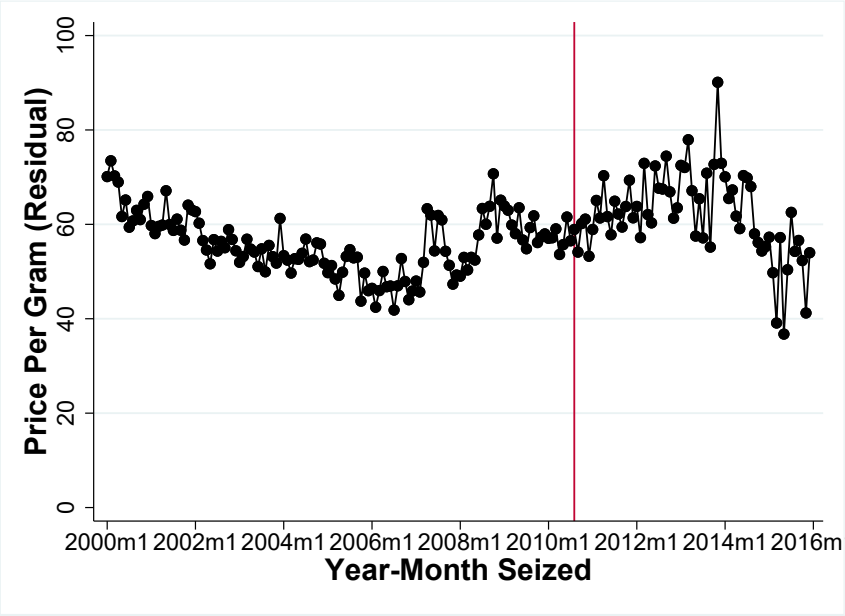
Notes. The figures above plot the distribution of charged amounts pre-2010 from 60-280g by race for various subsets of offenders. Panel (a) plots the distributions for offenders with above median criminal history scores. Panel (b) plots them for offenders with above median predicted sentenced (using exogenous factors, like age, sex, education, etc.). Panel (c) plots them for offenders convicted in states with above median levels of racial animus. In each plot, the distributions are similar. Kolomogorov-Smirnov tests of the equality of distributions fail to reject the null in all three cases: (a) p-value=0.54; (b) p-value=0.74; (c) p-value=0.98. This shows that the distributions of drugs charged from 60-280g are similar by race even within observably similar groups. These plots are created from the USSC data. See Table 1 for notes about data construction.

Figure A8. Additional Figures for Conditional Racial Disparity Tests



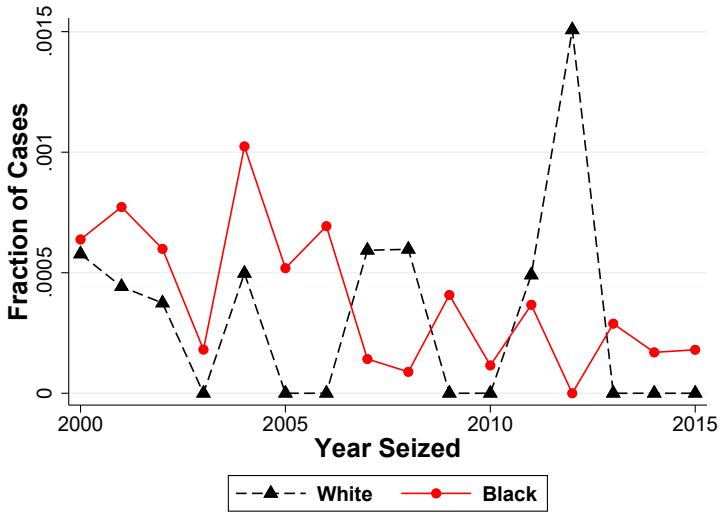
Notes. The figure in panel (a) plots the histograms of crack-cocaine amounts seized for white offenders and for black and Hispanic offenders from 0-10g. The white offenders are slightly over-represented at 1g, but otherwise, the distributions are very similar. The figure in panel (b) plots the histograms by race from 10-280g. White offenders are slightly over-represented at 20-30g, but otherwise, the distributions are very similar. These figures use the balanced sample of agencies (i.e. agencies that are present in all 16 years) in NIBRS. Panels (c) and (d) plot the coefficient δ^X from equation (4) of the main text for each bin starting at X divided by the share of cases in that bin (to calculate a percent difference). Since estimates are noisier at higher amounts, panel (c) shows the estimates for amounts from 60-380g in 20g bins. Panel (d) shows the full range of estimates for amounts from 60-1000g. Panel (e) plots the coefficient δ^X for each 10g bin from a regression that includes the standard controls and includes the interaction of the After2010 binary variable with a binary variable for above median criminal history and a binary variable for above median predicted sentence. This addresses concerns that the racial disparity in movement away from narrow ranges could be due to other differences even within those ranges. Panel (f) plots a similar figure except the estimates are from a regression which excludes controls entirely. Panel (g) plots the racial disparity in each bin from Figure 2b in red and overlays it with a plot of placebo estimates in black and gray. The placebo estimates are calculated from 100 replications in which I randomly assign white offenders to each 10g bin at the same rate as black and Hispanic offenders pre- and post-2010. The black circles plot the average of those placebo estimates and the gray bars plot a 90% confidence interval on those estimates. These plots are created from the USSC data. See Table 1 for notes about data construction.

Figure A9. Drug Prices Before and After the Fair Sentencing Act, DEA



Notes. This figure plots the drug price per gram (conditional on state, drug potency, type of drug, month seized, and a linear trend in year) against the year-month the drugs were seized. Outliers above the 95th percentile ($200 per gram) and below the 5th percentile ($20 per gram) are excluded. The price is smooth and increasing through the date the Fair Sentencing Act was implemented. In other words, there is no clear price response in the illegal drug market, at least in the short run. I formally estimate the discontinuity around the date the bill was signed using a bandwidth of +/- 24 months and various polynomials (linear, quadratic, cubic). The estimated discontinuity is never statistically different from zero, and it ranges from -5.5 to 2.1. This plot is created from the DEA STRIDE data.

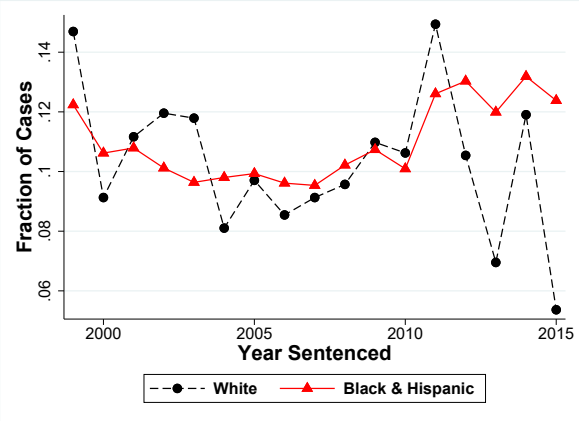
Figure A10. Fraction of Crack-Cocaine Seizures from 280-290g, Full Coverage States, NIBRS



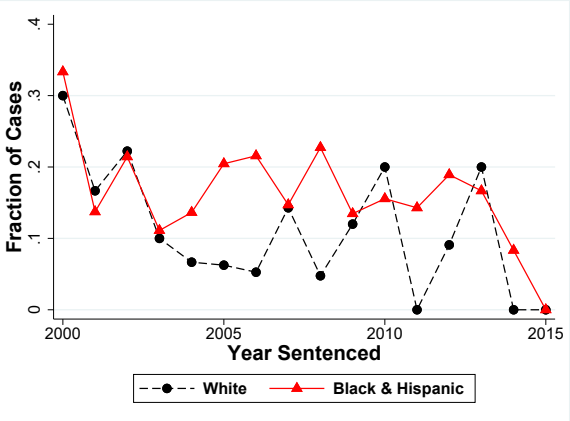
Notes. The figure above plots the fraction of crack-cocaine seizures with 280-290g by race. The sample is limited to states with full coverage (i.e. all agencies in the state participating) starting in 2012 and with 90% coverage or more from at least 2008-2015. This plot is created from the NIBRS data.

Figure A11. Alternative Figures Testing for Shifting from State/Local Authorities to Federal Court

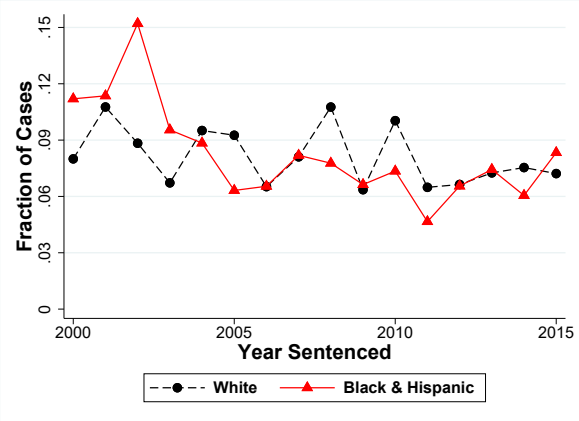
(a) Fraction of Cocaine Cases 200-400g, USSC



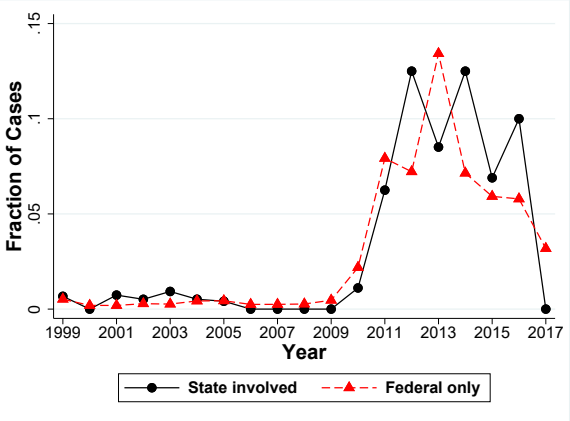
(b) Fraction of Cocaine Cases 200-400g, NC



(c) Fraction of Cocaine Cases 200-400g, FL, Alternative Sample



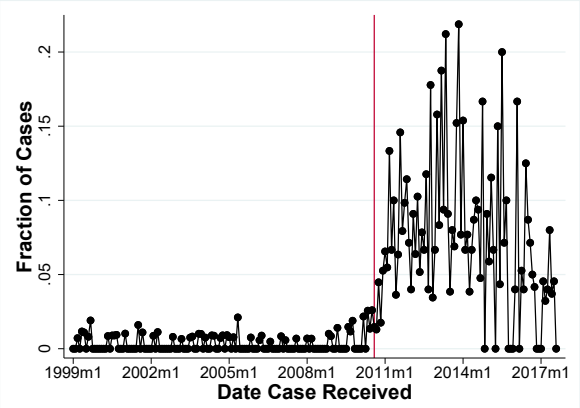
(d) Fraction of Crack-Cocaine Cases in 280-290g, by Type of Source Agency, EOUSA



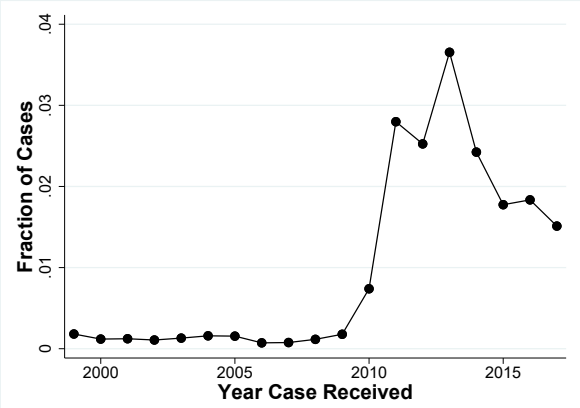
Notes. The figure in panel (a) plots the fraction of cocaine offenses with 200-400g in the USSC federal sentencing data, by race. The figure in panel (b) plots the fraction of cocaine offenses that have a range from 200-400g in NC state prison from 2000-2015, by race. Many of drug convictions in NC do not include type of drug in the offense description, the figure above is limited to those offenses that specifically list 'cocaine' in the offense description. The figure in panel (c) plots the fraction of cocaine offenses with 200-400g in FL state prison by race, limiting to those offenses that list a weight range in the offense description (the figure in the main text includes all cocaine offenses and codes (Convicted 200-400g)=0 if there is not a weight listed in the offense description). The figure in panel (d) plots the share of cases sent to EOUSA attorneys from sources that involve state agencies (red dashed line with triangle markers) and the share of cases sent to EOUSA attorneys from strictly Federal sources (black solid line with circle markers). This figure is limited to the top agencies sending cases and excludes joint investigations (e.g. FBI + state/local task force). The top agencies are: DEA, FBI, ATF, and state/local.

Figure A12. Fraction of Cases with 280-290g Over Time, EOUSA

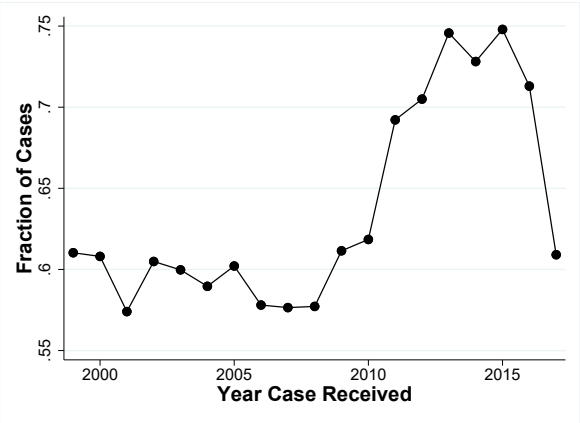
(a) By Month Received by EOUSA



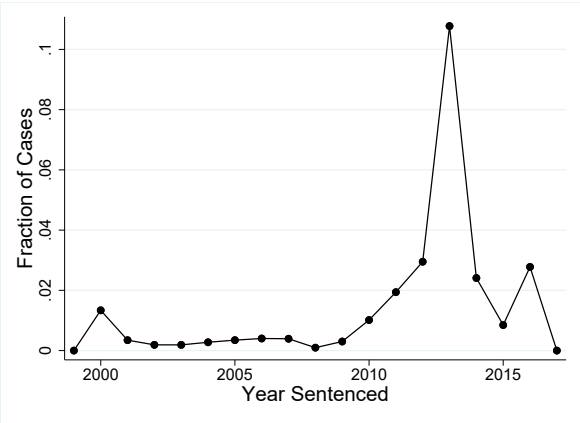
(b) Imputing Missing Weights as (280-290g)=0



(c) Imputing Missing Weights as (280-290g)=1



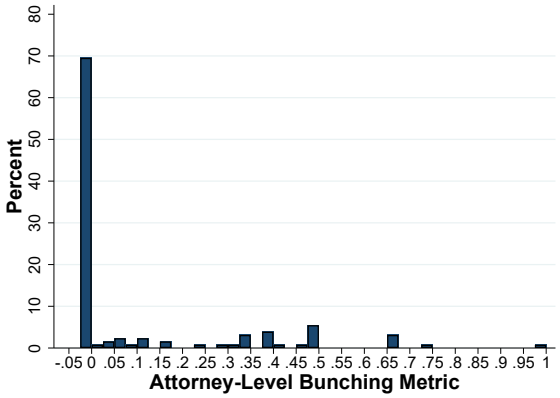
(d) Cases Received Before FSA



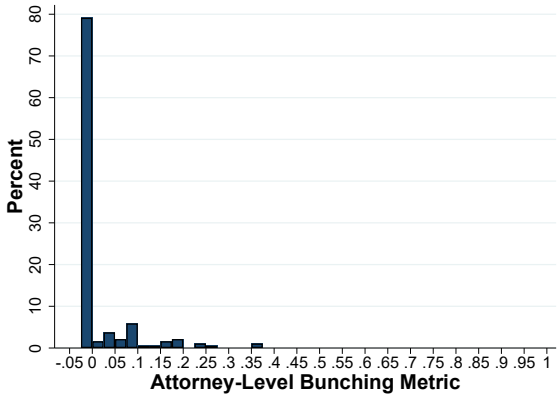
Notes. Panel (a) plots the fraction of cases with 280-290g (excluding cases with missing drug weights) by the month the case was received. The vertical red line indicates the date the Fair Sentencing Act was passed. In panel (b), I re-code the 280-290g dummy variable equal to zero if the drug weight is missing (typically, I leave the dummy variable missing if the drug weight is missing). In panel (c), I do the opposite, coding the 280-290g dummy variable equal to one if the drug weight is missing. In both cases, there is a sharp increase in the fraction of cases at 280-290g after 2010. Since panel (b) more accurately matches the statistics from the USSC final sentencing data, I use that imputed value for various robustness tests. Appendix E explores further robustness to alternative ways of dealing with missing values. Panel (d) plots the fraction of cases with 280-290g in each year for cases that are received by the EOUSA prior to the signing of the Fair Sentencing Act. These plots are created from the EOUSA data.

Figure A13. Histograms of Attorney-level Bunching Metric at 280-290g, EOUSA

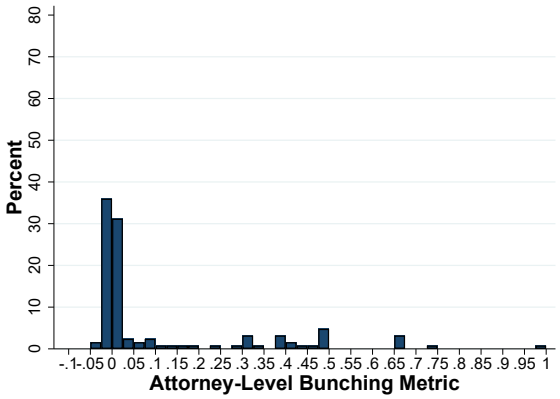
(a) 280-290g, National



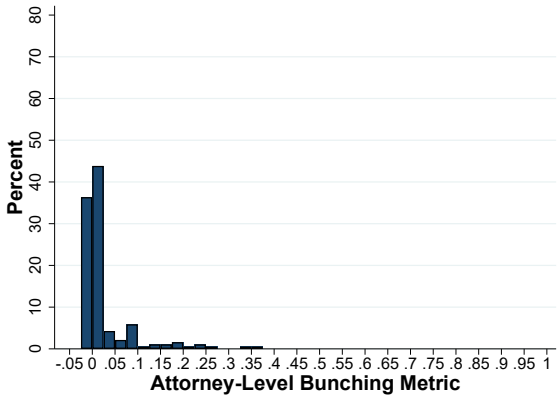
(b) 280-290g, Missing = 0, National



(c) 280-290g, District



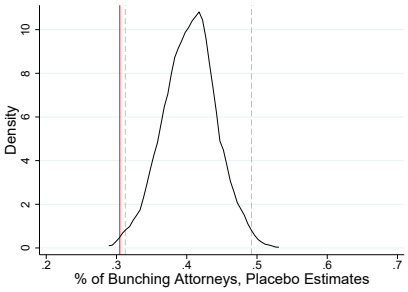
(d) 280-290g, Missing = 0, District



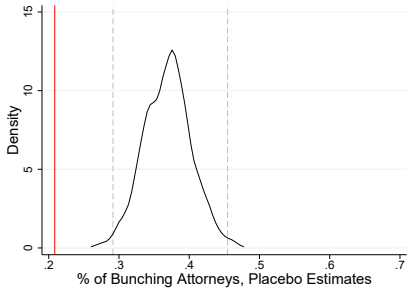
Notes. The figures above plot histograms of attorney-level bunching metrics, which are calculated as the difference between each attorney's fraction of cases with 280-290g post-2010 and the average fraction of cases with 280-290g at "baseline." In the national case (panels (a) and (b)), the baseline is the average fraction of cases with 280-290g prior to 2010. In the district case (panels (c) and (d)), the baseline for an attorney in district **A** is the average fraction of cases with 280-290g prior to 2010 in district **A**. Panels (b) and (d) include cases where the drug weight field is missing by coding the 280-290g dummy variable equal to zero when the drug weight is missing. I define an attorney as a "bunching attorney" if their bunching metric is above zero, thus the exact fraction of bunching attorneys for each panel is as follows: (a) 30.5%, (b) 20.9%, (c) 31.2%, and (d) 20.9%. These figures are limited to attorneys with 10+ cases post-2010. Limiting to 15+ cases delivers similar results: (a) 30.5%, (b) 22.9%, (c) 31.0%, and (d) 22.9%. Limiting to 5+ cases decreases the fraction of bunching attorneys to: (a) 21.2%, (b) 14.2%, (c) 21.4%, and (d) 14.2%. Even imputing missing weight cases as though they **are all** 280-290g cases (the highly unrealistic result in Figure A12c) implies that only 70% of attorneys bunch at 280-290g. These plots are created from the EOUSA data.

Figure A14. Histograms of Randomized Attorney Bunching Metric, EOUSA

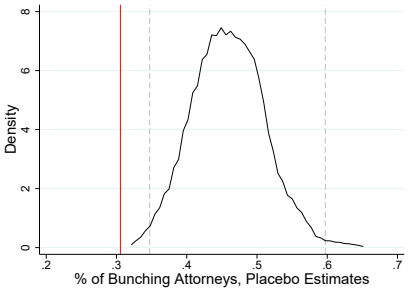
(a) 280-290g, 10+ cases



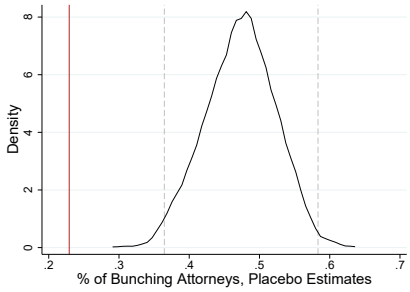
(b) 280-290g, Missing = 0, 10+ cases



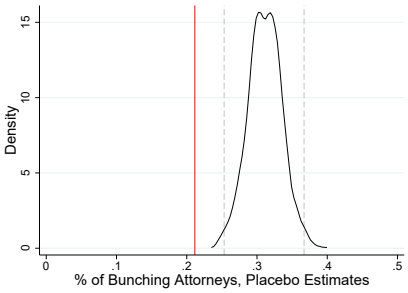
(c) 280-290g, 15+ cases



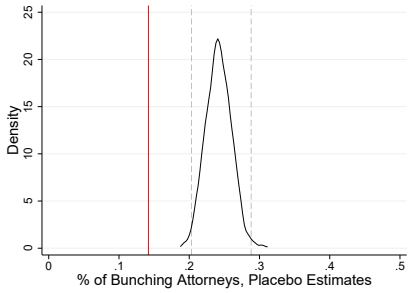
(d) 280-290g, Missing = 0, 15+ cases



(e) 280-290g, 5+ cases



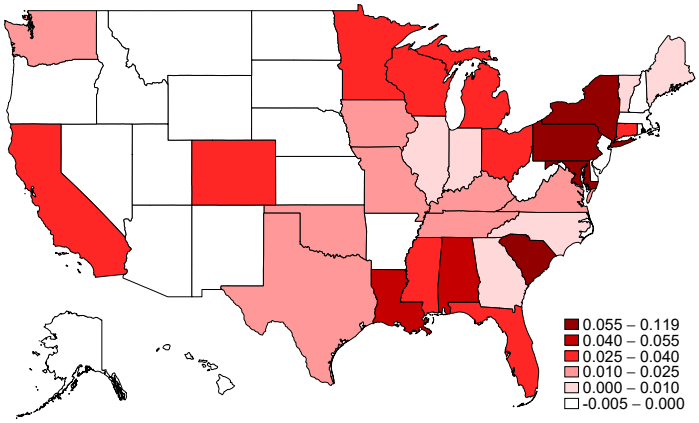
(f) 280-290g, Missing = 0, 15+ cases



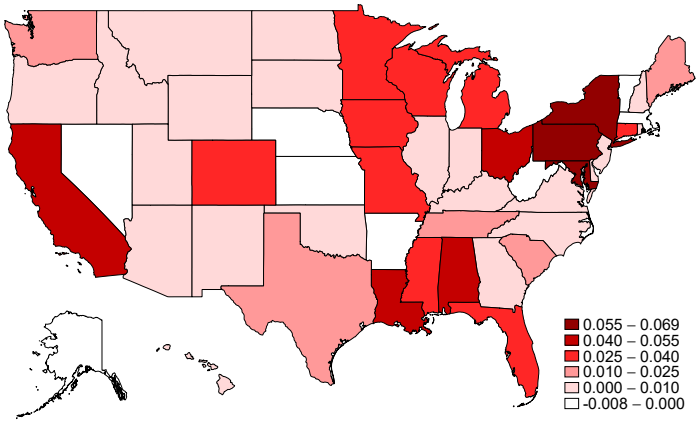
Notes. Please note the different axes in panels (e) and (f). In panel (a), I randomly re-assign all cases in the sample of attorneys with 10 or more cases after 2010, maintaining the same overall fraction of 280-290g cases in each year. After doing this random re-assignment, I calculate the number of bunching attorneys. I do this 1,000 times and plot the placebo estimates. Panel (b) does the same exercise but includes data with missing values imputed as zeroes. Panels (c)-(d) use attorneys with 15 or more cases and panels (e)-(f) use attorneys with 5 or more cases. The gray dashed lines indicate the 1st and 99th percentiles of the placebo distribution and the red line indicates the actual fraction of bunching attorneys from the data. These plots are created from the EOUSA data.

Figure A15. Map of State-level Bunching and State-level Racial Disparity in Bunching

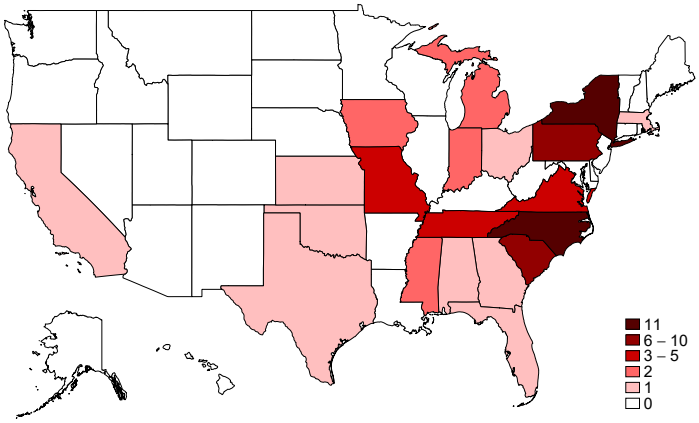
(a) Bunching Coefficient from USSC



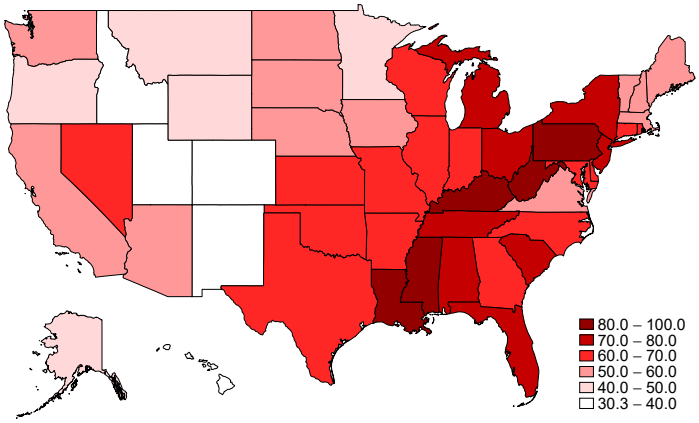
(b) Non-White/White Difference in Bunching from USSC



(c) Number of Bunching Attorneys (5+ Cases) from EOUSA



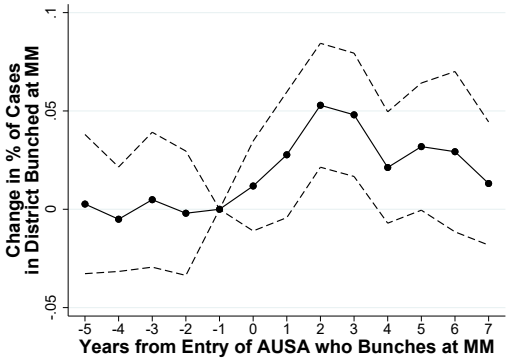
(d) Google Racial Animus Measure



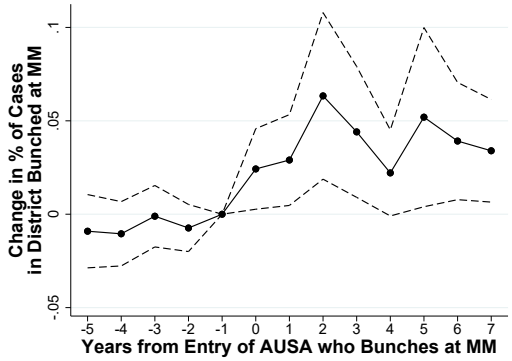
Notes. Panel (a) plots the state-level bunching estimate for all states with a sufficient number of cases. Panel (b) plots the difference between the state-level bunching estimate for white offenders and the state-level bunching estimate for black and Hispanic offenders for all states with a sufficient number of cases. These plots are created from the USSC data. See Table 1 for notes about data construction. Panel (c) plots the number of prosecutors who bunch in each state (among those prosecutors with 5+ drug cases after 2010). This plot is created from the EOUSA data. Panel (d) plots the racial animus index derived from Google search volume for a racial slur and introduced by Stephens-Davidowitz (2014). For Panels (a) and (b) there are several states that do not have enough cases to estimate bunching or racial disparities in bunching at 280-290g (these states are: AZ, DE, HI, ID, MT, ND, NH, NJ, NM, NV, OR, RI, SD, UT, WY). I pool all of these states in one regression and apply the resulting coefficient.

Figure A16. Additional Evidence of Prosecutorial Discretion in Bunching, Movers Results, EOUSA

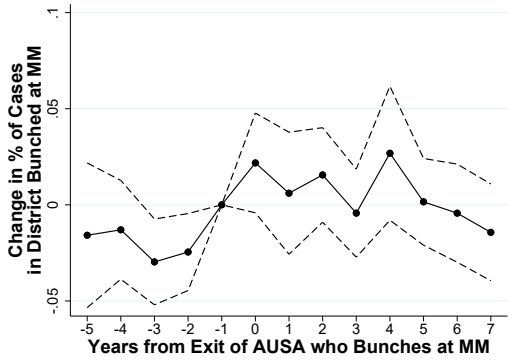
(a) Effect of Entry of a Bunching AUSA



(b) Effect of Entry of a Bunching AUSA, Low-Bunching Districts



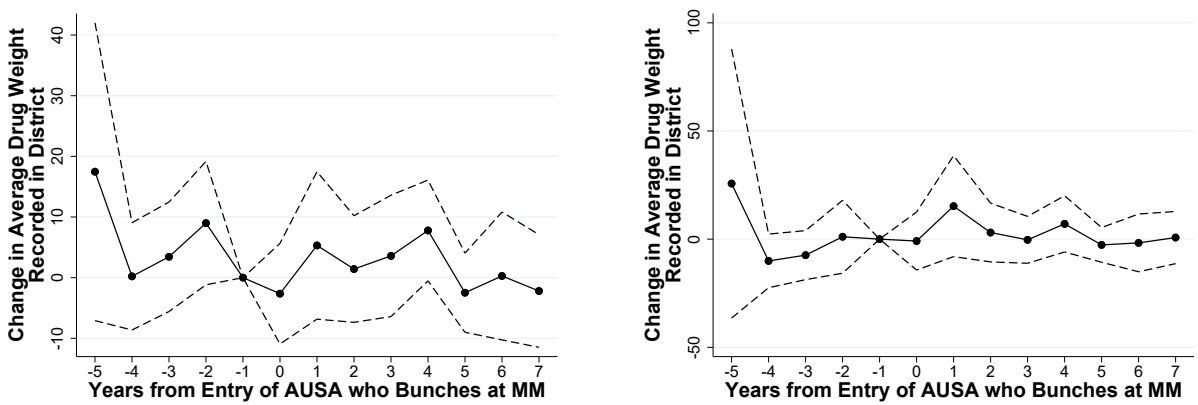
(c) Effect of Exit of a Bunching AUSA



Notes. Panels (a) and (b) plot the change in the percent of cases that are bunched at the mandatory minimum (MM) threshold (50g pre-2010 and 280g post-2010) after a "bunching" prosecutor enters a district. For these figures, I identify prosecutors who switch districts, who bunch at the mandatory minimum threshold in their first district, and who have 5 or more cases in their first district. I then identify the districts that they switch into and analyze the fraction of cases bunched at the mandatory minimum for all other prosecutors in that district. Panel (a) shows that prior to entry of a bunching prosecutor, district-level bunching does not change year-to-year, but that immediately after the bunching prosecutor enters, other prosecutors in that district increase their fraction of cases bunched at the threshold. Panel (b) shows that this increase is driven by districts that have low-levels of bunching (below the median for all districts) prior to the entry of the bunching prosecutor. Panel (c) plots the bunching activity for the districts from which these prosecutors are leaving. This analysis is limited to the first bunching attorney from panels (a) and (b) that leaves the district. There is not a decrease in the prevalence of bunching after bunching prosecutors exit a district. This suggests bunching at the mandatory minimum threshold is not related to a temporary behavior shift, such as increased competition among attorneys, but that it may be related to something more permanent, such as learning about techniques or developing beliefs/norms. The dashed lines in panels (a)-(c) are 90% confidence intervals. Since these figures rely on prosecutors who move from one district to another and require reasonably long pre- and post-periods, I use data from 1994-2016 and identify the first moving attorney for post-1999 years only (insuring a 5-year pre-period for every district). In practice, this means the figures above are largely based on bunching at 50-60g (the pre-2010 mandatory minimum). Restricting to post-2010 moves does not yield a large enough sample of movers with sufficient cases to classify them as bunching versus non-bunching. Also, since these figures involve cross-district analysis and missingness varies across districts, I impute missing weights as zero instead of excluding them from the analysis. The results are robust but noisier when excluding missing weights (the post-period increase is 0.029, p-value=.18 for all districts and is 0.098, p-value=0.002 for low-bunching districts. See Appendix E for more detail). These plots are created from the EOUSA data.

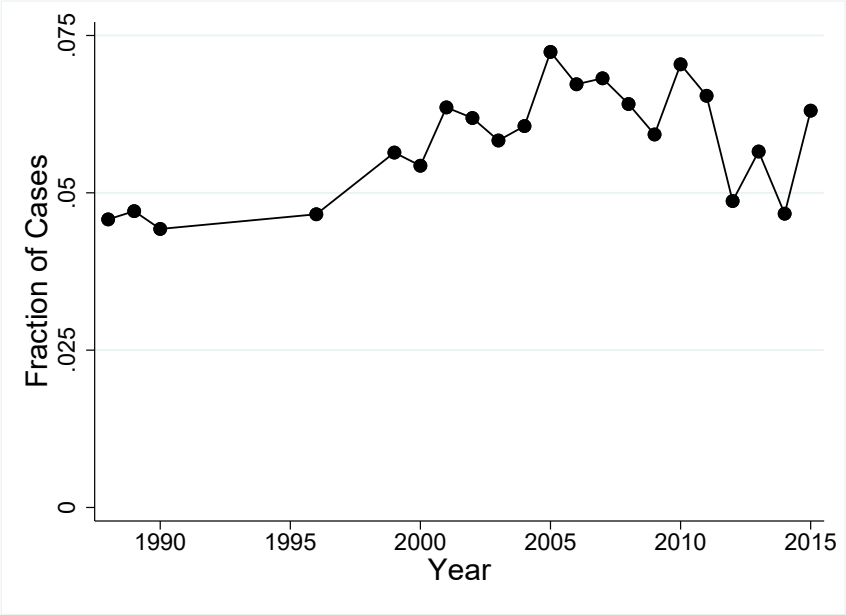
Figure A16. Additional Evidence of Prosecutorial Discretion in Bunching, Movers Results, EOUSA

(d) Drug Weight **(e) Drug Weight, Low-Bunching Districts**



Notes. Panels (d) and (e) plot the change in average drug weights recorded after a "bunching" prosecutor enters a district. For these figures, I identify prosecutors who switch districts, who bunch at the mandatory minimum threshold in their first district, and who have 5 or more cases in their first district. I then identify the districts that they switch into and analyze the average weight recorded for all other prosecutors in that district. One concern with Figures A16a-c is that bunching prosecutors may be brought into a new district because the composition of cases that district receives is changing. However, both panels show that drug weights, on average, are not increasing before or after the prosecutor enters the district. The dashed lines in panels (d)-(e) are 90% confidence intervals. See Figures A16a-c for additional notes about data construction. These plots are created from the EOUSA data.

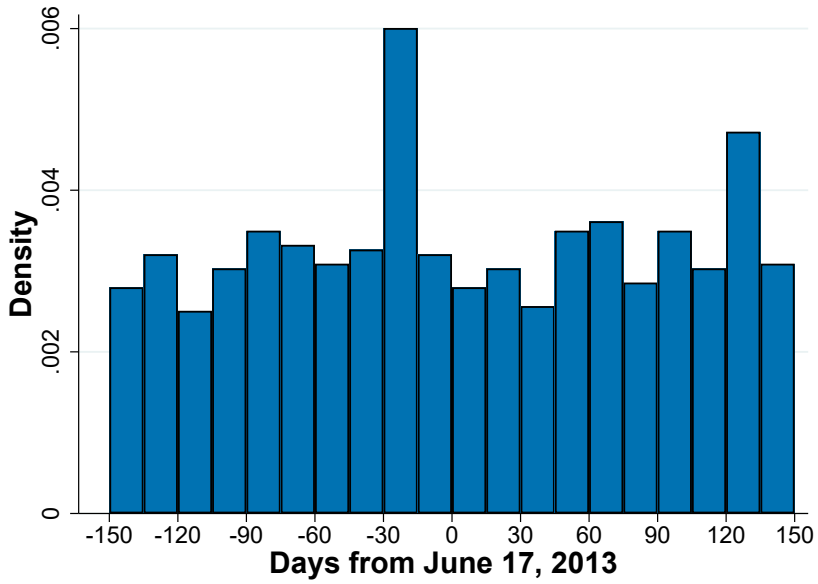
Figure A17. Fraction of Cases in 50-60g by Year, from USSC Sentencing Data



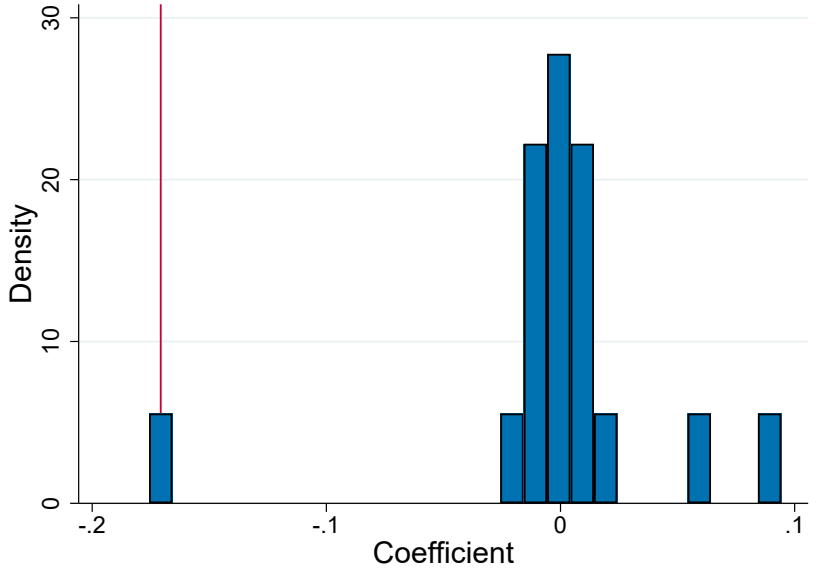
Notes. The figure above plots the fraction of all cocaine (powder and crack) cases with 50-60g by year. The sample is limited to cases with drug weights from 0-1000g. All cocaine cases are used because earlier years (1988-1990) do not distinguish between types of cocaine. This figure indicates that cases bunched above the pre-2010 10-year mandatory minimum threshold increased by about 60% from 1988-90 to 2010. Over this same time period, the average weight of cases from 0-1000g decreased. This suggests that the practice of bunching cases at the mandatory minimum was potentially learned over time, which is consistent with the evidence on movers and the spread of bunching in Figure A16. This plot is created from the USSC data.

Figure A18. Tests of Validity for *Alleyne v. US* Result, EOUSA

(a) Density of Cases Received Around June 17, 2013 (Date of Decision in *Alleyne*)



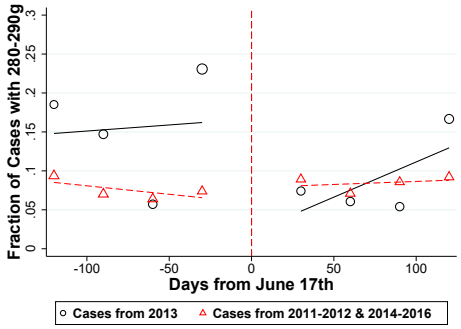
(b) Estimate of Discontinuity Around June 17 in All Years 1999-2016



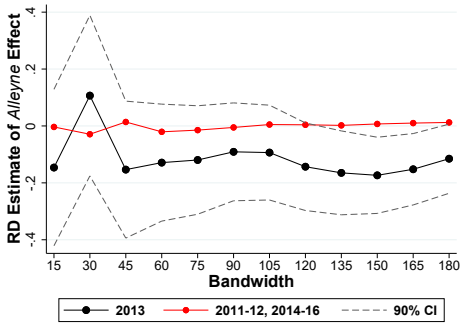
Notes. Panel (a) plots the density of cases around the June 17, 2013 (centered at zero) and grouped into 15-day bins. June 17, 2013 is the day *Alleyne v. US* was decided. Outside of the large number of cases from -30 to -15 days before *Alleyne* was decided, the density is relatively smooth through that date. Panel (b) plots a histogram of the estimated discontinuity around June 17 in all years from 1999-2016. The estimates are centered at zero and the coefficient in June 2013 (marked by the red line) is twice as large as the next largest estimate of any sign and over 4 times larger than the next largest negative estimate. These plots are created from the EOUSA data.

Figure A19. Robustness of *Alleyne v. US* Result to Choice of Bandwidth and Polynomial, EOUSA

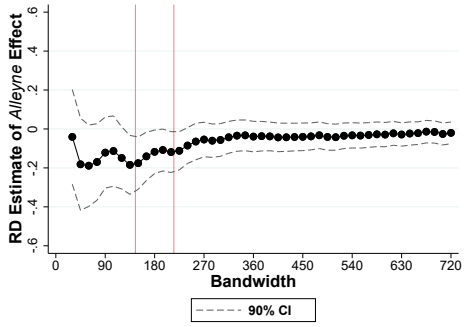
(a) Main Result, Change in Bunching after *Alleyne*



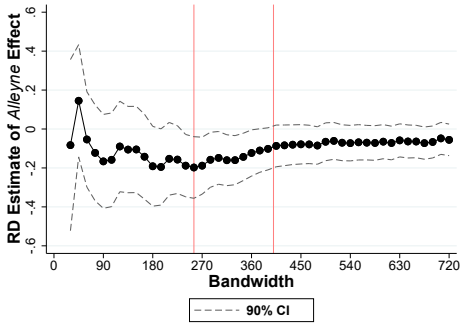
(b) RD DiD, Linear



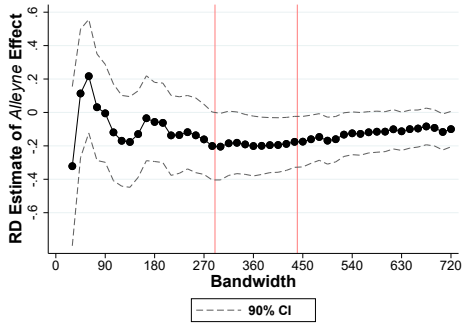
(c) RD, Linear



(d) RD, Quadratic



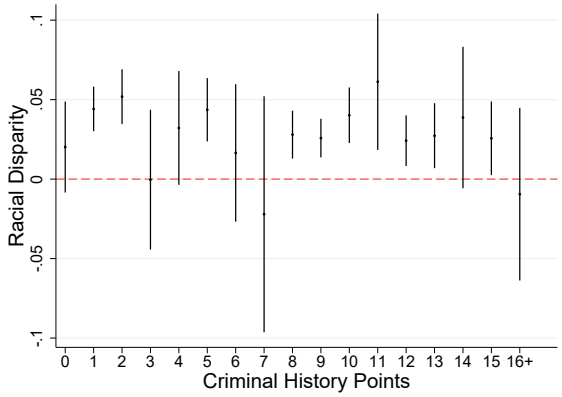
(e) RD, Cubic



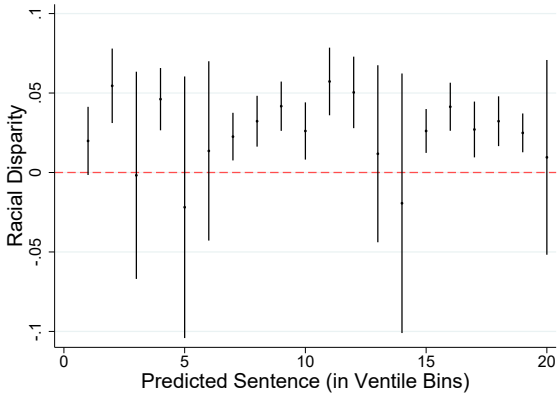
Notes. The figures above display estimates for the effect of *Alleyne v. US* (a case that strengthened evidentiary requirements) on the prevalence of bunching at 280-290g. Panel (a) plots the fraction of cases with 280-290g in each 30-day bin for 120 days before and 120 days after June 17th. The black circles show the fraction of cases in each bin for 2013 and the red triangles show the average fraction of cases in each bin for 2011-2012 and 2014-2016. The solid black line shows a linear fit on each side of the June 17, 2013 and the dashed red line shows a linear fit on each side of June 17 for all other years. The scatter plot symbols are weighted by the total number of cases in each bin. The estimated discontinuity is $\delta = -0.1433$ and $se = 0.0935$. Panels (b)-(e) display estimates across many different bandwidth choices (i.e. the number of days before and after June 17 included in the regression) and different polynomial choices (i.e. the polynomial of the running variable, number of days from June 17, included in the regression) are shown across panels. Panel (b) displays coefficient estimates from the RD difference-in-differences regression for bandwidths from 15-180. Since the difference-in-difference estimates use multiple years, bandwidths above 160 days are asymmetric. The black line in panel (b) displays the estimates from 2013, the red line displays the estimates from all other years after 2010 (when nothing in particular happened around June 17). Panels (c)-(e) estimate a typical RD regression (i.e. not using variation around June 17 in other years). This allows me to extend the bandwidth to 2 years before and after *Alleyne v. US*. In these panels, the first red line denotes the CER-optimal bandwidth and the second red line denotes the MSE-optimal bandwidth (Calonico et al. 2017). In panel (c), for example, the estimate approaches zero at larger bandwidths–this is to be expected. As we get further from the cutoff, the linear polynomial becomes an increasingly bad fit. In all three panels, the optimal bandwidths yield estimates that are statistically different from zero (or marginally statistically significant). These plots are created from the EOUSA data.

Figure A20. Heterogeneity in Disparity by Detailed Breakdown of Criminal History and Predicted Sentence

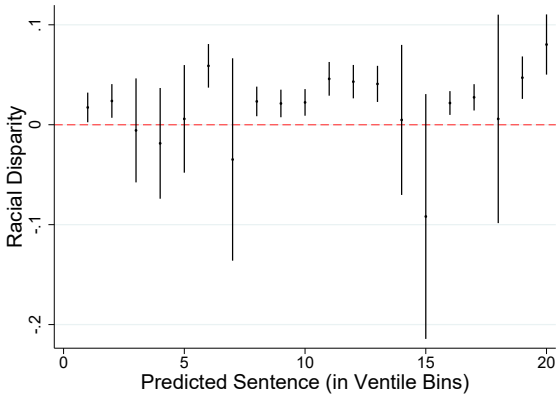
(a) Criminal History



(b) Predicted Sentence (Exog)

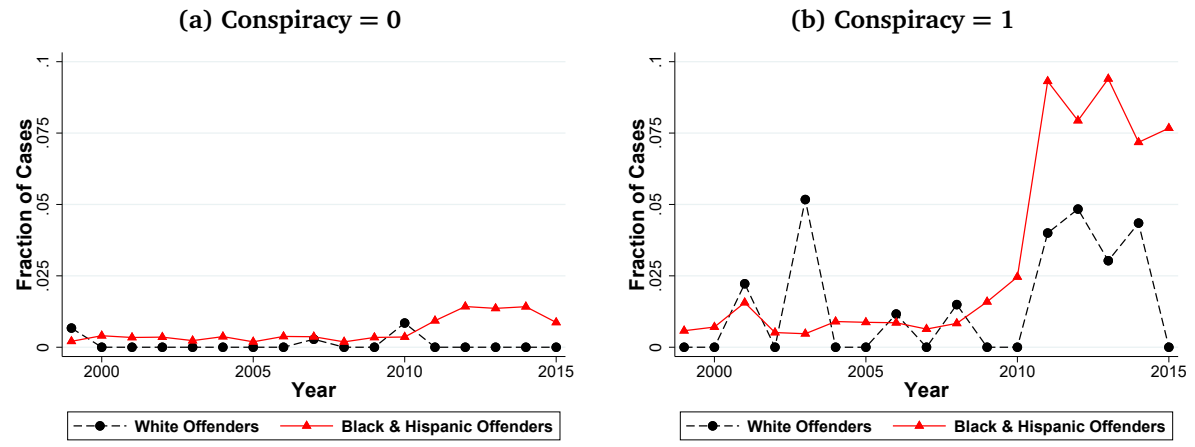


(c) Predicted Sentence (Endog)



Notes. The figures above display the estimated racial disparity in bunching for various subsets of offenders based on criminal history score, predicted sentenced (based on exogenous factors), and predicted sentenced (based on exogenous and endogenous factors). Panel (a) estimates the racial disparity within criminal history score, grouping scores of 16+ together (about 5% of the sample). This is similar to Figure 4b but looks at criminal history points individually instead of grouping them together. The mean sentences for criminal history scores of 0, 5, 10, and 15 are: 5.6 years, 8.8 years, 11.6 years, and 12.6 years. Panel (b) estimates the racial disparity within ventiles of predicted sentence (based on exogenous factors: criminal history score, sex, citizenship status, age, number of dependents, education, and district). This is similar to Figure 4c but looks at ventiles instead of quantiles. The mean sentences for ventiles 1, 5, 10, 15, and 20 are: 4.1 years, 6.4 years, 9.0 years, 11.0 years, and 12.7 years. Panel (c) estimates the racial disparity within ventiles of predicted sentence (based on exogenous and endogenous factors: the factors listed for panel (b) plus number of current counts, statutes involved in the case, whether there was only one drug involved, whether crack-cocaine is the primary drug, and whether the career offender, aggravated role, or violent offender enhancements were applied). The mean sentences for ventiles 1, 5, 10, 15, and 20 are: 3.8 years, 6.0 years, 7.8 years, 10.6 years, and 18.4 years. The correlation between this predicted sentence and actual sentence post-2010 is 0.6. These plots are created from the USSC data. See Table 1 for notes about data construction.

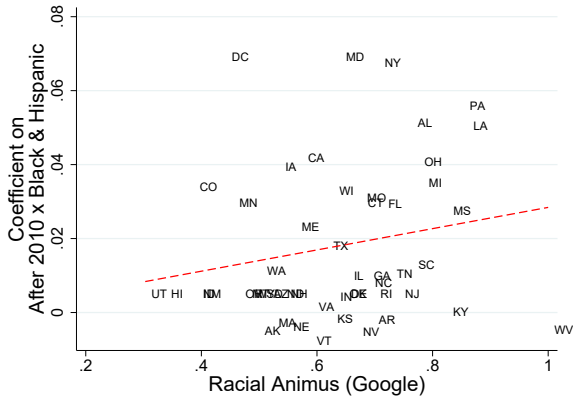
Figure A21. Heterogeneity in Disparity by Presence of a Conspiracy Charge



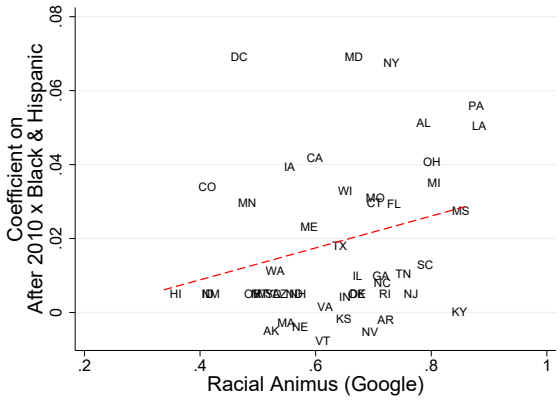
Notes. Panel (a) displays the fraction of cases recorded with 280-290g in each year by race for cases that do not have a "conspiracy" charge. As expected, bunching is much lower in these cases since conspiracy charges are one mechanism used to increase the amount charged beyond the physical amount seized. However, there is still a statistically significant racial disparity in bunching in these cases (see Table 9). Panel (b) displays the fraction of cases recorded with 280-290g in each year by race for cases that have a "conspiracy" charge. As expected, bunching is much higher in these cases. Again, there is a statistically significant racial disparity in bunching in these cases (see Table 9). These plots are created from the USSC data. See Table 1 for notes about data construction.

Figure A22. Relationship between State-level Disparity and State-level Animus

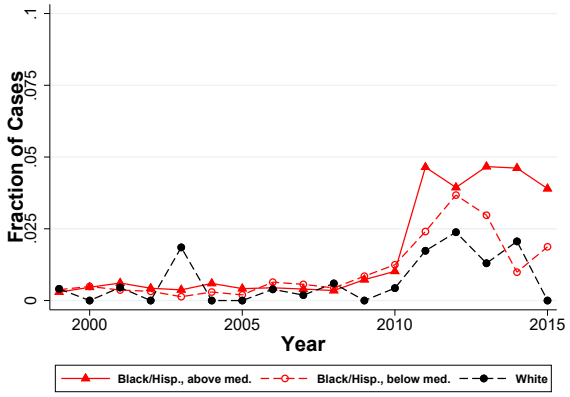
(a) All States, with Shrinkage



(b) Excluding Outliers, with Shrinkage



(c) Heterogeneity in Disparity by State-level Animus Over Time



Notes. Panels (a) and (b) plot the racial disparity in bunching estimated at the state level. There are several states that do not have enough cases to estimate racial disparities in bunching at 280-290g (these states are: AZ, DE, HI, ID, MT, ND, NH, NJ, NM, NV, OR, RI, SD, UT, WY). I pool all of these states in one regression and apply the resulting coefficient. Since individual states do not have many cases, the state-level estimates are sometimes very large in magnitude but very noisy. I adjust these state-level estimates using a Bayesian shrinkage procedure that shrinks each estimate, based on its standard error, to the mean of the distribution (Chandra et al. 2016). Panel (a) includes all states and panel (b) excludes states with racial animus above the 99th percentile or below the 1st percentile. The coefficient estimate on the linear fit in panel (a) is 0.029 (p-value=0.19). The coefficient estimate on the linear fit in panel (b) is 0.043 (p-value=0.07). The coefficient estimate from regressing the estimates with shrinkage on a dummy variable for above median animus is 0.010 (p-value=0.08). Panel (c) displays the fraction of cases recorded with 280-290g in each year for black and Hispanic offenders convicted in states with above median levels of racial animus (red solid line with triangle markers); black and Hispanic offenders convicted in states with below median levels of racial animus (red dashed line with open circle markers); and white offenders (black dashed line with closed circle markers). See Table 8 and Table 9 for formal estimates of this relationship. These plots are created from the USSC data. See Table 1 for notes about data construction.

Appendix B. Additional Institutional Details

Numerous articles in law journals and from criminology and related fields discuss the existence of prosecutorial discretion in US courts. The idea that prosecutors have wide latitude over how to handle cases is not controversial. Ulmer et al. (2007) concludes, "prosecutors have great influence through charging, sentence bargaining, and, in the case examined here, the application of mandatory minimums." In this paper, I study how prosecutors exercise that discretion after the FSA raises the mandatory minimum thresholds for crack-cocaine offenses. Through a series of tests, I find evidence that prosecutors increase amounts charged in response to the FSA; specifically, they increase the number of cases charged right above the 10-year mandatory minimum threshold. I also find that the burden of this discretion falls on black and Hispanic offenders. This is consistent with prior work that also documents a racial disparity in the use of discretion in federal courts (e.g. Fischman and Schanzenbach 2012; Rehavi and Starr 2014).

Despite this concordance with prior work, some may remain concerned about the specific setting in this paper. I address many concerns briefly in the main text of the paper, but I discuss them in more detail and provide additional institutional context below.

First, can prosecutors argue the amount involved in the case is greater than the physical amount seized? Yes, the USSC Guidelines (2015) state, "Where there is no drug seizure or the amount seized does not reflect the scale of the offense, the court shall approximate the quantity of the controlled substance." Quantity is first recorded in the Presentence Report (PSR), which is prepared by a probation officer who takes input from the prosecutor and the defendant (Webb and Turow 1982). While it is possible for probation officers to depart from the prosecution's account of events in the PSR, anecdotal evidence suggests the prosecution's input is critical. In *Hard Bargains*, Mona Lynch quotes a defense attorney who explains that the probation officer's determination depends on what the prosecutor turns over to them, "what happens is the probation officer then requests a copy of the prosecutor's file. So if you have a prosecutor who is on board with making sure this is how it's going to go, they will only turn over to the probation office [what's] in their file that supports that recommendation." Even if the PSR does differ from the prosecution's account, the prosecution has the opportunity to object and argue their case for a lower amount to the judge. Empirically, I observe bunching at 280g in the EOUSA case management records, which are recorded before the PSR occurs.

Once the report is prepared, the prosecution and defense are given the opportunity to make objections to the report. Before *Alleyne*, the judge determines the facts at sentencing based on a preponderance of evidence standard (i.e. is it more than 50% likely that the claim is true?) and records the quantity used at sentencing in the Statement of Reasons (SOR). Based on USSC data, the quantity recorded in the SOR rarely differs from the quantity recorded in the PSR. After *Alleyne*, drug quantities that trigger the mandatory minimum must be proven to a jury beyond a reasonable doubt (i.e. is it more than 99% likely that the claim is true?). I discuss *Alleyne* in more detail below.

Lynch also notes that drug weight is frequently subject to negotiation between the prosecution and defense. It is clear from the accounts in her book that prosecutors are able to argue the amount involved is greater than the amount seized. She describes two approaches used to increase the quantity. First, prosecutors can argue that the defendant was involved in a conspiracy and that the weight attributable to the conspiracy as a whole should be attributable to each defendant. Second, they can introduce testimony from informants about the defendant's drug involvement, including type and quantity of drugs sold. She emphasizes that this weight, "does not require anyone in law enforcement actually to have even seen any drugs of those or other quantities to count against the defendant. It can be established through hearsay or other questionable means and still be counted as relevant conduct."

She explains that prosecutors can use these approaches to induce pleas by threatening to have informants testify about more deals or the scale of the conspiracy. She quotes a defense attorney who explains how this discretion affects the decision to plead guilty or go to trial, "If you're dealing with a larger-scale conspiracy, then the government can conceivably add drug weight [...] Depending on how many informants they have that can say, for instance, your client knew about X amount of drugs or was involved with different transactions dating back, you know, to other periods of time. It just is a very ambiguous, dangerous place to be if you're thinking about taking it all the way in terms of a trial." In one interview with Lynch, an AUSA describes this process exactly, "The actual heroin sales directly tied to Mr. Samuels and his son were of 1g and 4g, respectively; the rest was arrived at on the mere say-so of confidential informants. [...] She told me that she could have established enough historical weight, through those (conspirators)

she had 'flipped,' to get Mr. Samuels to at least a ten-year mandatory minimum sentence, if not more."

Fischman and Schanzenbach (2012) describe this phenomenon as well, "these minimums often permit discretion in their application. For example, drug quantity is often an issue of serious contention in large conspiracies or when trafficking has occurred over time. [...] The prosecution exercises considerable discretion over whether to attempt to prove or charge quantities that would invoke the mandatory minimum."

There are other ways that prosecutors can argue for higher quantities. Drug cases may rely on wiretapped phone calls or text messages as evidence of wrongdoing. In those records, the defendants may use code that must be translated to an actual weight, and discretionary choices must be made as to how code is translated. Lynch notes that sometimes informants take on this role of translating slang or estimating weights based on photos or videos.

Attorney Dan Honold, writing in the Harvard Journal on Legislation, describes two additional methods. First, he notes the importance of conspiracy charges for increasing the weight involved in the case. He also explains that in addition to conspiracy charges, the prosecution can argue that money seized as part of the arrest should be used to estimate drug quantity. Cash can be converted into quantities by using an estimated sale price per unit of drugs and dividing the amount of money by that price. Finally, he describes how weight can also depend on how the drug is weighed and whether other non-drug substances are counted. The weight used in the quantity calculation could include both the weight of the drug and the weight of the "carrier medium."

Honold also notes that this limits judicial discretion by constraining judges' decisions with mandatory minimums. He acknowledges that these rules that allow prosecutors to argue larger quantities may sound fair or intuitive but, ultimately, have unfortunate consequences, "at first glance these [...] seem to be clarifying rules by which to calculate an unclear quantity of controlled substances in a given case. In effect, however, they provide many opportunities for the quantity calculation to be inflated at the sentencing phase, leading to higher sentences."

These are not abstract arguments about what prosecutors could feasibly do under the law, but rather descriptions of approaches that they frequently use. I discuss some examples from actual court cases below.

Second, why do prosecutors argue for larger quantities than the amount seized and why would they do this disproportionately by race? Prosecutors argue for quantities that trigger mandatory minimums for two main reasons: (1) they can be used to induce plea deals and cooperation with the government and (2) they can increase sentences and limit judicial discretion at sentencing. There are numerous reasons why prosecutors may value these outcomes. For example, ending a case in a plea requires fewer resources than going to trial, avoids the uncertainty of trial, and may lead to the defendant's cooperation in other cases. However, even in cases that end in a plea, there can be disputes over the amount attributed to the defendant at sentencing, with the defense arguing for a lower amount and the prosecution arguing for a higher amount. At this stage, the plea or the trial has already taken place, leaving sentence length as the primary concern. Prosecutors may value long sentences for career reasons or because they believe they are effective as a deterrent for future crime or as a just punishment for the crime that was committed.

Although it is not possible to know each prosecutor's motive for seeking plea deals or long sentences, we do know that at least a small fraction of prosecutors engage in explicit misconduct to achieve those things. Focusing on lawyers in general, Rozema (2020) finds that 4.4% of lawyers who pass the bar exam are publicly sanctioned within the first 25 years since licensing. In Federal court, there have been 44 exonerations since 2000 that are in part due to prosecutorial misconduct (National Registry of Exonerations). To be clear, arguing that a defendant was involved with a quantity that is greater than the amount seized is legally permissible and would not necessarily qualify as misconduct. However, it does require the prosecutor to exercise discretion and it is important to understand whether they use that discretion fairly (i.e. quantities are accurately attributed to defendants) and equitably (i.e. prosecutors don't disproportionately exercise lenience for one race).

Below, I discuss some empirical and qualitative evidence about whether the discretion is applied fairly and equitably. First, I detail reasons why this discretion might be applied inequitably. In the main text of the paper, I discuss and test several explanations for why black and Hispanic offenders are more likely to be charged right above the mandatory minimum threshold after 2010 than white offenders. One candidate explanation is that black and Hispanic offenders have more drug involvement, and the bunching at 280g reflects this greater drug involvement. I show evidence that this is not the case. The main piece of evidence is that prior to 2010, black and Hispanic offenders have a very similar quantity distribution from 60-280g compared to white offenders, but that after 2010, black and Hispanic offenders are less likely to be charged in narrow ranges from 60-280g (e.g. 160-170g, 190-200g, etc.).

The pre-2010 distributions suggest that these offenders have similar drug involvement, and the post-2010 shifts suggest that despite this similar involvement, black and Hispanic offenders are now less likely to be charged with those quantities and more likely to be charged with quantities that trigger the mandatory minimum.

Another candidate explanation is that black and Hispanic offenders may be different from white offenders in other ways, and that those differences lead to the disparate treatment. I also provide evidence which suggests this is not the case. For example, observable characteristics like criminal history or education do not explain the racial differences in bunching at 280g. A third candidate explanation is that the "cost" of bunching is lower for black and Hispanic offenders. For example, black and Hispanic offenders may have access to worse defense counsel or judges might view weak evidence more favorably in the case of black and Hispanic offenders. I show evidence that the bunching at 280g is not related to judges or defense counsel, but ultimately, the courtroom setting is complex, making it hard to disentangle prosecutor actions or beliefs from those of other actors.

A fourth candidate explanation is that prosecutors discriminate against black and Hispanic offenders. This could be a result of statistical discrimination, implicit or unconscious racial animus, or explicit racial animus. It is not possible to completely disentangle these three explanations. In the paper, I show that the racial disparity in bunching at 280g is correlated with a state-level measure of racial animus. Is it feasible that prosecutor decisions might be influenced by racial animus, whether explicit or implicit? To show that this is feasible, I highlight the experiences of defendants and examples of prosecutors (not limited to federal prosecutors) saying or doing things that reveal biased or racist attitudes. The following five paragraphs contain descriptions of racist behavior and some relevant quotes (slurs omitted). Again, I include these examples and quotes to highlight the experience of defendants and to display direct, narrative accounts of racial bias from prosecutors.

The Supreme Court recently overturned the conviction of Curtis Flowers due to racial bias in jury selection (*Flowers v. Mississippi*). The Court writes, "A review of the history of the State's peremptory strikes in Flowers' first four trials strongly supports the conclusion that the State's use of peremptory strikes in Flowers' sixth trial was motivated in substantial part by discriminatory intent. [...] In the six trials combined, the State struck 41 of the 42 black prospective jurors it could have struck."

In an article on racially biased language in prosecutorial summations, Praatika Prasad documents several glaring examples. For one, "a Black man from South Carolina, was convicted of murder and sentenced to death in 2000. [...] During trial, the prosecutor [...] referred to [him] as a 'monster,' 'caveman,' and a 'beast of burden,' and he unnecessarily brought up [his] sexual history with a 'blond-headed lady.' Later, in his closing argument, [the prosecutor] referred to [him] as 'King Kong.' The court overruled defense counsel's multiple objections, and the all-white jury sentenced [him] to death. Six years later, one of the jurors [...] stated that he believed that [he] had killed the alleged victim '[b]ecause [he] was just a dumb [slur].' [...] Finally, in 2016, a federal trial judge in South Carolina overturned [his] death sentence after recognizing that the trial 'was so infected by racial animus by the prosecutor and a juror . . . that [he] was deprived of his constitutional right to due process." Prasad uses this case to highlight that one can develop race-neutral explanations for behavior that is, in fact, racially biased. He notes that various courts interpreted the prosecutor's actions as race-neutral for sixteen years until a judge ultimately recognized that it "stoked race-based fears."

Throughout the article Prasad notes several other instances of racially biased speech from prosecutors including: prosecutors referring to black defendants as monkey characters from TV or movies; a prosecutor using a quote from an enslaved person in the movie Gone with the Wind to imply that a black witness would lie for a black defendant; prosecutors emphasizing to the jury that a black defendant committed the crime in "our streets" and not "some ghetto"; and a prosecutor quoting "Dixie", a song adopted by Confederate secessionists.[1]

The Chicago Tribune documents another example of racism by prosecutors (Possley and Armstrong 1999). In the article, a former prosecutor describes the culture of the county's prosecutors office, "There was an ongoing competition among prosecutors to be the first to convict defendants whose weight totaled 4,000 pounds. [...] Because most of the defendants were African-American, Goggin recalls now, with no small degree of discomfort, the competition was described in less sensitive terms behind closed doors–'[slur]s by the Pound.'"

Lu (2007) introduces qualitative evidence from a focus group of twelve former US Attorneys, many of whom recognized the existence of racial disparities in criminal justice and the role of prosecutor discretion and bias. One

[1] In the case of the quote from "Dixie", the court explicitly noted it, "references with praise a time and place for the most pernicious racism"

former US attorney described a scenario with an AUSA being particularly lenient for a white defendant, "I had an [AUSA who] wanted to drop the gun charge against the defendant [in a case in which] there were no extenuating circumstances. [...] [the AUSA] said, 'He is a rural guy who grew up on a farm. The gun he had with him was a rifle. He is a good ol' boy, and all the good ol' boys have rifles, and it's not like he was a gun-toting drug dealer.' But he [was] a gun-toting drug dealer, exactly." Another former US attorney describes being the first Hispanic AUSA hired in a border district and hearing their colleague refer to defendants with an ethnic slur.

These examples do not imply that **every** prosecutor is acting on racial bias. As the evidence from the focus group demonstrates, some prosecutors or former prosecutors are aware of these biases and are working to eliminate them. Ultimately, it is not clear how prevalent these attitudes are among prosecutors, but the examples above show that it is at least feasible that some prosecutors may have these beliefs, either consciously or unconsciously, and that they may act on them. The GSS provides some information about the racial attitudes of lawyers more generally. Among the small sample of people reporting an occupation of "lawyer": 20% oppose (or strongly oppose) a close relative marrying a black person; almost 10% rate black people as more unintelligent than intelligent; 10% rate black people as more lazy than hardworking; 25% say racial differences are due to a lack of will among black people; 71% oppose busing; 70% oppose affirmative action; and 42% say that black people should work their way up without special favors. However, very few respondents support interracial marriage laws or report being unwilling to vote for a black president, two common measures of explicit racism. Notably, 40% say courts are not harsh enough in dealing with criminals whereas only 20% say they are too harsh. These statistics are based on all years in the GSS (1972-2018), but the percentages are similar when limiting to more recent years (2000-2018).

Third, are the argued amounts still lower than true involvement; in other words, are the argued amounts based on evidence that could have been used to prove the defendant was involved with an even greater amount? One explanation for the bunching at 280g is that prosecutors are conservatively estimating quantity and that they could produce evidence of further drug involvement, if necessary. This is a difficult claim to test directly because I don't observe the same information that the prosecution observes. However, I provide evidence that cases bunched at 280g are based on relatively weak evidence. First, I find that bunching decreases after a Supreme Court case that increases the evidentiary standard in mandatory minimum cases. I discuss this case and its implications in more detail below. Second, although I am not able to link sentencing data to appeals data at the case-level, I do find that districts with more bunching at 280g are also more likely to have drug cases reversed or remanded on appeal and specifically more likely to have them reversed or remanded for an issue with drug quantity. In addition to this empirical evidence, I discuss examples from Federal cases that involve quantity calculation.

Lynch interviews a defense attorney who casts doubt on informant testimony, "In one of the transactions which was recorded, my client made a comment about, you know, 'Oh, you know, this stuff is so popular I'm selling X amount a week.' Or 'X amount at a time.' And [the prosecutor] was arguing that that quantity should be used to up the weight even though there was no stuff there. My argument was that he was trying to placate a dissatisfied customer by telling him that the quality was so good. You know, just puffing." Another defense attorney also comments on the quality of the evidence in these cases, "She said that in the state system, where she previously worked, drugs had to actually be possessed to charge with intent to sell, 'but here we have this historical weight, and you know, trying to explain to a client how the law allows another individual to throw out numbers—'I've known this person for a year. They've been my source of supply for a year. I would purchase this substance in these quantities. And I did that for six months.' . . . It blew my mind when I came to the federal system. What do you mean, drugs that they weren't caught with, and based upon all of these snitches' statements?'" Finally, Lynch highlights the incentive issues embedded in this system, noting that informants often benefit from their statements against other defendants making the testimony, "especially unreliable as evidence."

Now I discuss several court cases where these issues are brought to the fore. It is not possible to link cases in the USSC data to court documents because the data do not include case numbers. Also, the primary search tool for court documents, PACER, does not allow a keyword search and can only be searched by case number or defendant name or by filtering broadly on district or date. This makes finding cases with quantity calculation issues difficult. I use RECAP and judyrecords, searchable databases of a subset of PACER documents, to find example cases. To be clear, these cases may be cases in which the defendant legitimately was involved with the alleged amount. I include these cases not to allege wrongdoing but to illustrate instances where discretion played a role in establishing quantity.

In *US v. Eric Lanard Williams*, the quantity attributed to the defendant hinged on the government's interpretation of slang used to describe drug amounts. The court writes, "He [the defendant] explained that the government misinterpreted certain code words that he used during drug negotiations which were captured on audio recordings and in text messages. Specifically, he claimed that when he used the terms two of the usual, one, or a whole one, he was referring to a quarter-ounce or seven grams of crack cocaine, as opposed to an entire ounce."

In another case, *US v. Hector Morales*, the quantity calculation was based on an assumption about how much powder cocaine the defendant cooked into crack-cocaine. This is a critical assumption because triggering the 10-year mandatory minimum requires 5000g of powder cocaine as opposed to 280g of crack-cocaine. The case also highlights the use of informants to establish quantities. The court writes, "The evidence at trial included testimony from a cooperating government witness regarding Morales personal involvement with 150 grams of crack, as well as audio recording of Morales himself bragging to a government informant about his many drug sales, with quantity specifications ranging from 100, 200 grams as an estimate of his usual sale in Plattsburgh, to a specific statement of having sold 90 grams just prior to one phone conversation. These quantities referred only to grams and did not distinguish whether Morales was referring to powder or crack cocaine, but the specific evidence regarding the 150 grams of crack indicated that his modus operandi was to cook half the powder cocaine into crack."

US v. Andre Johnson, which was decided when the 10-year threshold was 50g, offers an example of a case in which the drug quantity included the weight of the drug and the weight of non-drug substances. The court describes the issue, "The drugs weighed 53.3 grams at the time of Mr. Johnsons arrest and weighed 33.15 grams at the time of the initial TBI lab report. Ms. Flowers, a forensic chemist with the TBI, testified that the disparity in weight was reasonable because over three months had elapsed between the arrest and the testing so that 40% quantity of water or solvents could have evaporated causing the crack cocaine to weigh less. The weight of the drugs is determined at the time of the offense and will include any solvents used to make the drugs"

In *US v. Kizzy Fader*, the petitioner appeals an earlier decision in which the government relied on witness testimony to estimate the weight of drugs involved in a conspiracy. The court explains, "Petitioner, in coordination with other members of the conspiracy, and with the conspiracy's knowledge, entered the property and recovered a quantity of crack cocaine hidden in an oven door, which quantity amounted to 35 bundles. [...] Additional evidence, including witness testimony, supported a finding that 35 bundles of crack cocaine would weigh 280 grams."

In *US v. David Tyrone Murrell*, the government initially argues for a quantity of 418g, by adding the 168g the defendant physically possessed to 250g based on testimony from an informant. The judge rejects the evidence for the 250g, which suggests that in this case, the evidence that did not even meet the preponderance of evidence standard. The government then revises the PSR and uses the approach of converting seized cash to an estimated drug quantity. The appeals court describes this, "The revised PSR, however, converted $6,000.50 in cash that Murrell possessed at the time of his arrest on March 9, 2011, into cocaine base (crack). Thus, the total drug weight in the revised PSR was 295.58 grams, and included the 168 grams of cocaine base (crack) that Murrell possessed on March 9, 2011, when he was arrested, and 127.58 grams of cocaine base (crack) from the conversion calculation arising from the $6,000.50." This calculation implies an estimated price per gram of $47.03. If the government had assumed a slightly higher price per gram of $54, the quantity attributed to the defendant would be slightly less than 280g rather than slightly above it.[2] In that case, the AUSA also explicitly states why they are seeking a longer sentence, namely that the defendant was "dishonest" at previous sentencing hearings. The defendant, on the other hand, explains their version of events, "What I'm guilty of is possession of 168 grams of crack, and accepting responsibility for it, to try to get back to my family and my kids."

In each of these cases, the government introduces evidence on drug quantity that might result in an inaccurate depiction of the offense. At least, this evidence is often less verifiable than evidence on amounts seized. In one example, a judge explicitly states the evidence submitted does not meet the burden of proof (*US v. David Tyrone Murrell*). The results from *Alleyne* and from the USSC appeals data further suggest that cases bunched right above mandatory minimum thresholds are based on relatively weak evidence.

The above discussion is concerned with whether prosecutors use these approaches to increase drug quantity fairly. A separate and important question is whether they use these approaches equitably. Even if prosecutors are accurately and fairly attributing quantities to defendants when they use these approaches, they may still be more

[2]Note, this also assumes that all of the cash seized was earned by selling crack-cocaine. At the assumed price of $47.03, the defendant would be responsible for a quantity under 280g if as little as 13% was earned by other means (even other drug sales).

lenient for one race than another. The empirical evidence in this paper suggests that prosecutors are perhaps unfairly harsh on black and Hispanic defendants but at least inequitably lenient for white defendants. Prior to 2010, the distributions of quantities charged from 60-280g are very similar by race. After 2010, black and Hispanic offenders are disproportionately less likely to be charged in the 60-280g range and disproportionately more likely to be charged in the 280-290g range. Even if those 280-290g charges are accurate, this raises the question: why aren't white offenders shifted to the 280-290g range at the same rate?

For example, in the case of David Tyrone Murrell, the defendant was held responsible for 295.58 grams of crack-cocaine. In the USSC data, one white offender is held responsible for 278.42g of crack-cocaine. These are just two examples, but on average, white offenders are more likely to be held responsible for 60-280g after 2010 relative to black and Hispanic offenders. Prosecutors may be more likely to rely solely on physical evidence in cases for white defendants or less likely to use non-physical evidence to inflate the weight above the threshold. In *US v. Dana Cooper*, the judge acknowledges that the prosecution may charge lower amounts as a form of lenience.

Fourth, why don't judges recognize and stop this practice? First, only 3.5% of cases are bunched at 280-290g after 2010. In practice, any given judge will only see a few of those cases in a year. Second, if those cases bunched at 280-290g are accurate but the bunching is inequitably applied, that would be a particularly difficult pattern for a single judge to detect. Third, until 2013, judges must determine the facts of the case based on the preponderance of evidence standard. This is a relatively weak legal standard that means it is more than 50% likely that the claim is true. Even if the judge finds the evidence questionable in a given case, it may still satisfy the preponderance of evidence standard. In 2013, a Supreme Court decision changes the legal standard for evidence in mandatory minimum cases. Below, I discuss a case in which a judge explicitly acknowledges that evidence might have met a lower legal standard but does not meet the new, beyond a reasonable doubt requirement. Still, judges do occasionally depart from the quantity attributed to the defendant in the PSR. In the example from *US v. David Tyrone Murrell*, the judge decided that the informant's testimony to 250g did not meet the preponderance of evidence standard.

Fifth, is examining bunching at the mandatory minimum threshold a good test for this use of discretion? I discuss this briefly in the main text of the paper. Conceptually, the prosecutor is charging the minimum amount necessary to trigger the mandatory minimum in these cases, suggesting that they may be based on weaker evidence than cases charged with amounts far above 280g. Empirically, there is a substantial increase in bunching at 280g, but only small increases above 290g after 2010. After accounting for trends in drug weights charged, those small increases above 290g are the same magnitude for black, Hispanic, and white offenders. This means that the racial disparity in whether a case is sentenced above 280g at all is the same magnitude, albeit statistically noisier, as the racial disparity in whether the case is bunched narrowly above 280g.

Interestingly, researchers at the Bureau of Justice Statistics proposed a similar test in a 2015 working paper. They describe the motivation as follows, "Drug offenses provide one venue for finding evidence of whether prosecutors have exercised discretion to the disadvantage of blacks. Mandatory minimums for drug violators are triggered by the amount of drugs that were trafficked. [...] We look for evidence of prosecutorial manipulation in recorded drug weights for drug trafficking cases. Specifically, we look to see whether blacks are more likely than whites to be above a mandatory minimum threshold for drug cases, with amounts near a threshold that triggers the application of mandatory minimum sentencing laws. If blacks are systematically disadvantaged, then blacks should be more likely than whites to be above a mandatory minimum threshold." I discuss that paper in more detail in the main text, but I highlight their motivation here as further support that measuring bunching above the 10-year threshold is a feasible method for detecting prosecutorial discretion and racial disparities in the use of that discretion.

Sixth, did *Alleyne* affect case outcomes? In *Alleyne v. US*, the Supreme Court decided that any fact that increases the statutory minimum for an offense is an element of the offense and must be decided by a jury beyond a reasonable doubt. Prior to *Alleyne*, facts about quantity were decided by a judge at sentencing based on a preponderance of evidence standard. The Supreme Court has long recognized that this distinction is critical. In *Jones v. US*, they write, "Much turns on the determination that a fact is an element of an offense rather than a sentencing consideration, given that elements must be charged in the indictment, submitted to a jury, and proven by the Government beyond a reasonable doubt." Lynch also acknowledges this issue pre-*Alleyne*, noting that drug weight did not need to proven

beyond a reasonable doubt and defendants thus had little ability to challenge it.

Empirically, I find that the fraction of cases bunched at 280g decreases after *Alleyne*, suggesting that those cases would not have withstood the scrutiny of a jury and the new legal standard. There is also qualitative evidence from court cases about the effect of the new legal standard on case outcomes. In *US v. Curtis Lee Dale*, a case decided after *Alleyne*, the court ignores the jury's finding with regard to weight and adopts a higher weight for sentencing purposes. The evidence the jury rejected and the court accepted were two empty wrappers that police found in a duffel bag, one of which the government estimated would hold 1kg of crack-cocaine. The defendant's petition to the Supreme Court argues, "The jury after a full trial on the matter outright rejected the government's contention for a higher quantity. It's absolutely inappropriate for a Sentencing Court to overrule a jury without ample evidence in the record which contradicts the jury verdict on quantity. [...] Dale's sentence should be vacated and remanded back to the Eighth Circuit Court of Appeals in accordance with *Alleyne* and *O'Neil*." *O'Neil* is a case that was decided after *Alleyne*. The jury decided on one quantity but the sentencing court adopted a different quantity. The Supreme Court vacated the judgment and remanded the case to the 8th Circuit Court of Appeals.

Finally, in an appeal of *US v. Marvin Lee Ellis*, the government explicitly states that evidence would have met a lower legal standard but did not satisfy the beyond a reasonable doubt standard required by *Alleyne*. The court writes, "The government points to the controlled buys from Tatum and Ellis totaling 25.3 grams. But that still leaves the government needing to show overwhelming evidence of another 254.7 grams needed to sustain the district court's sentence. [...] To show that 280 grams of crack cocaine are individually attributable to Ellis, the government relies primarily on the testimony of Ellis's nephew, Theoplis. Theoplis responded affirmatively to the government's question whether the powder cocaine Ellis purchased was "always cooked into crack cocaine. But Theoplis admitted that he had just once seen Ellis cook powder cocaine into crack. [...] On appeal based on Theoplis's testimony, the government argues that Ellis and Tatum cooked all their purchased powder cocaine into crack. If so, the government would have a strong basis to argue that Ellis and Tatum conspired to manufacture and distribute at least 280 grams of crack cocaine. But we are reviewing for constitutional harmless error, not for sufficiency of the evidence. And Theoplis's testimony does not provide the "uncontested and overwhelming evidence" necessary for us to find the *Alleyne* error harmless beyond a reasonable doubt."

Seventh, do mandatory minimums affect sentencing? In Appendix C, I show that cases sentenced just above the 280g threshold have longer sentences than cases sentenced just below it. This is particularly true in later years as the USSC Guidelines become more lenient. Is it feasible, however, that judges feel constrained by mandatory minimums such that they actually affect sentencing? Consider an example from the case *US v. Karen Antoinette Johnson*. In this case, the judge specifically noted that they wish to hand down a shorter sentence but that they were limited by the mandatory minimum. In an appeal, the reviewing court describes the issue, "In arriving at her offense level, the district court found Johnson was responsible for 361.86 grams of cocaine base [...] Were it free to impose a sentence it thought appropriate, the district court indicated it would vary downward from the advisory guidelines range and impose a sentence of less than eighty-seven months imprisonment. Because, however, it was bound by the provisions of 841(b)(1)(A), it imposed a sentence of 120 months imprisonment."

Last, it is possible that a person's sentence can fall below 10 years even when a mandatory minimum is triggered. This typically occurs when the offender receives a safety valve or substantial assistance departure, both of which require cooperation with the government. This cooperation can be costly for defendants, potentially even more costly than extra years on their sentence. See Table A9 notes for more detail. Since both cooperation with the government and increased sentences can be costly, I don't make a distinction between offenders with and without departures in the main analysis.

Appendix C. Sentencing Consequences

In order to understand the policy implications of bunching at 280g, I estimate the sentencing consequences of crossing the mandatory minimum threshold. Since mandatory minimums only constrain the minimum sentence, it is possible that being above the amount has no effect on actual sentencing.[3] I investigate this by employing an empirical strategy similar to Bjerk (2017) and estimating the following:

$$Sentence_i = \alpha + \beta_1 Above280_i + \beta_2 Amount_i + \beta_3 (Above280 \times Amount)_i + \epsilon_i \quad (1)$$

where $Sentence_i$ is the sentence handed down for offender **i**, $Above280_i$ is equal to one if the offender is recorded with 280g or more of crack-cocaine and zero otherwise, and $Amount_i$ is equal to the offender's recorded drug quantity centered at 280g. For the main results, I focus on cases sentenced after 2010. In Table C1, I estimate similar regressions using the pre-2010 data. As long as the offenders who are bunched above the threshold are not negatively selected from the population just below the threshold, then β_1 will provide a conservative estimate of the sentencing penalty associated with crossing the mandatory minimum threshold after 2010. The bunching above 280g suggests this assumption may be violated. As such, I also estimate (1) for states with low levels of bunching above 280g.

I find that bunching at 280g does have sentencing consequences. Offenders recorded with 270-280g after 2010 have a mean sentence of 9.6 years whereas offenders recorded with 280-290g after 2010 have a mean sentence of 11.4 years. Figure C1a plots sentencing outcomes by drug weight from 230-330g and the linear fit on each side of the 280g threshold for cases sentenced after 2010.[4] The discontinuity (β_1) is the sentencing penalty from crossing the mandatory minimum threshold. Figure C1b shows that there is no discontinuity in predicted sentence, where sentence is predicted from a model using pre-2010 cases and several offender characteristics. Figure C1c plots actual sentence for the subset of cases sentenced in states that have low levels of bunching. Even in states where there is little manipulation around the threshold, there is a sentencing penalty of about 2 years.[5,6] This result is robust to using the natural log of years sentenced and to using a binary indicator of whether the sentence was above 10 years (see Table C1). See Figure C1g for robustness to bandwidths from 10g to 250g.

Bjerk (2017), using data from fiscal years 2011-2012, employs a similar design and finds no sentencing penalty. I can fully replicate those results by limiting to those two fiscal years. The gist of the differences is that the USSC guidelines have changed over time, making sentencing for drug offenses below 280g more lenient and thus, giving the mandatory minimum more bite. Bjerk (2017) makes the important point that the sentencing effect of the mandatory minimum is related to the overall federal sentencing structure that crack-cocaine offenders face, and this exercise confirms that point. See Table C2 notes for a more detailed discussion of Bjerk (2017), the estimates of the sentencing penalty by year, and the reasons for variation in the penalty over time.[7]

The sentencing estimate in this Appendix assumes that an offender bunched at 280g would be charged with an amount just below 280g in the absence of the 280g threshold. However, the results in Section V.A.2 suggest that offenders bunched at 280g come from throughout the distribution below 280g. The average sentence after 2010 for offenders in the 50-280g range is 7.9 years. Using that value as the counterfactual sentence implies a sentencing consequence of 3.7 years.

[3] Judges could choose to treat defendants with 270g the same as defendants with 280g and sentence both to 10 years.

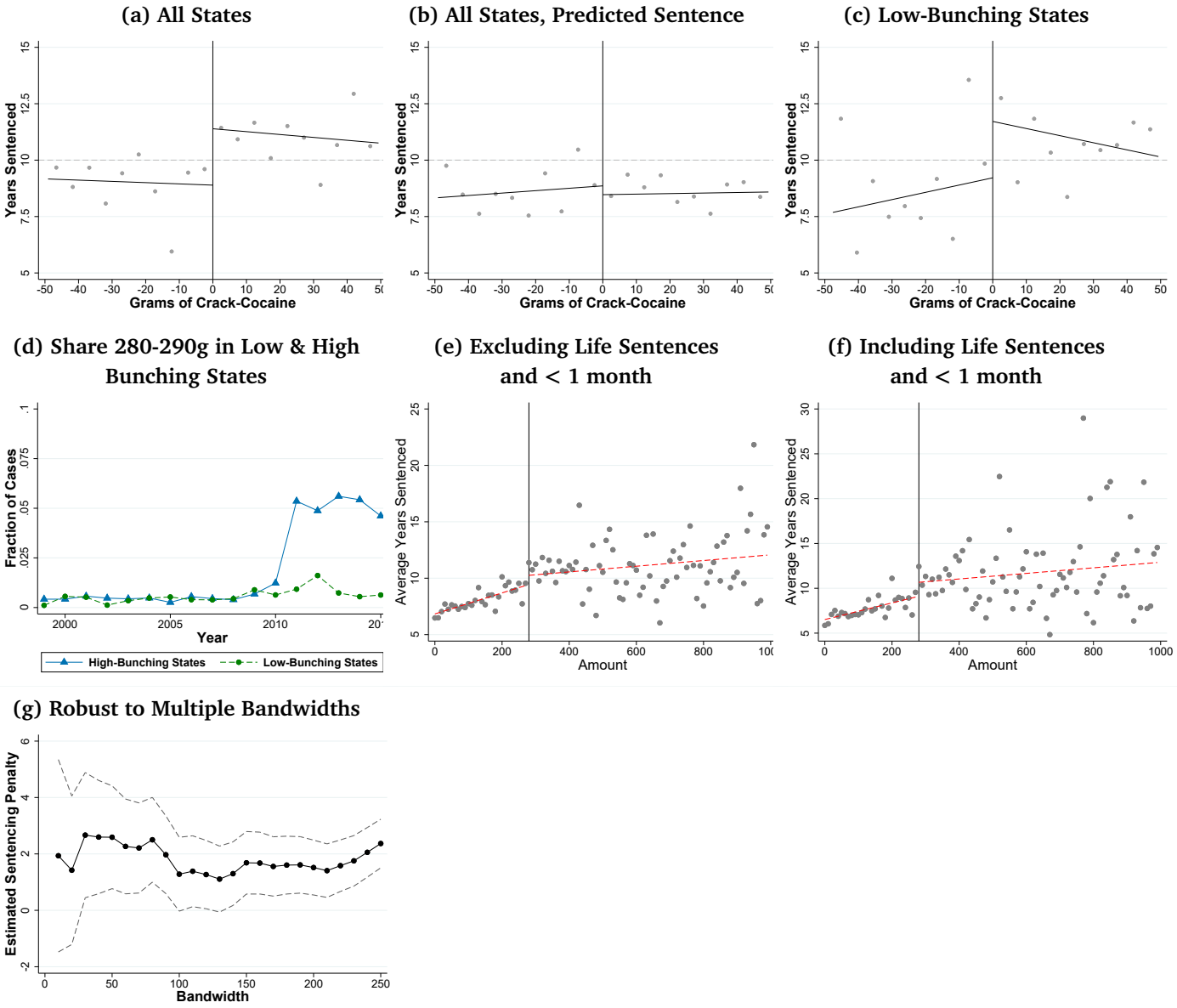
[4] Figures C1e-f plot sentence length across the full sample of drug weights from 0-1000g after 2010.

[5] This is possible because although offenders are negatively selected (in terms of sentence) on some characteristics, like race, they are positively selected on others, like criminal history score.

[6] These estimates indicate that there is a sentencing penalty for crossing the mandatory minimum threshold before and after 2010, not that sentences were longer or that the penalty of triggering the mandatory minimum was higher after 2010.

[7] Didwania (2020) shows that sentences exhibit small decreases when federal prosecutors are instructed to stop charging mandatory minimum quantities. Craigie and Zapryanova (2020) find sentences decrease overall after state-level mandatory minimum reforms. My analysis shows that when mandatory minimums can be used to increase sentences, they are effective in doing so. As such, the racially disparate application of mandatory minimums remains a serious policy concern.

Figure C1. Sentencing Consequences of Crossing the Mandatory Minimum Threshold



Notes. Figure C1a plots the average sentence (within each 5g bin) from 230-330g for cases sentenced after 2010. A linear fit is estimated on each side of the 280g threshold. The estimated sentencing discontinuity is about 2.5 years ($se = 0.89$). Figure C1b is the same plot but using predicted sentence from a model of sentencing and offender characteristics using pre-2010 data. There is no discontinuity in this figure, suggesting that offenders bunched at 280g are not negatively selected on characteristics that would increase sentence length in the absence of the threshold. Figure C1c is the same plot but limited to the subset of states that have low-levels of bunching. The estimated discontinuity is about 2.5 years ($se = 1.83$). These figures exclude life sentences and sentences less than 1 month, but results are robust to their inclusion (see Table C1). Figure C1d plots the share of cases with 280-290g by year for low- and high-bunching states. The coefficients described above are estimated from the regression:

$$Sentence = \alpha_0 + \beta_1 \text{Amount}_i + \delta_1 \text{Above280}_i + \phi_1(\text{Amount} \times \text{Above280}_i) + \epsilon_i$$

δ_1 is the estimated discontinuity (reported in the preceding notes) in sentencing due to crossing the mandatory minimum threshold. Figure C1e plots the relationship between years sentenced (excluding life sentences and sentences less than one month) and the amount charged in the case for cases sentenced after 2010. Figure C1f plots the same relationship but includes life sentenced (coded as 70 years) and sentences less than one month (coded as 0 years). Both figures demonstrate that even above 280g, sentences tend to increase in length slightly as the amount involved increases. Figure C1g plots the sentencing penalty of crossing the 280g mandatory minimum threshold after 2010, as estimated using the RD difference-in-difference model specified in Table C1 notes. The dashed lines are 90% confidence intervals. Estimates using a quadratic in polynomial are similar in magnitude but slightly noisier. The bandwidths used in the figure above range from 10g to 250g, in 10g intervals. These figures are based on the USSC data. See Table 1 for notes on data construction.

Table C1. Sentencing Consequences of Being Above the Threshold Amount

	Years Sentenced						Pr(Sentence ≥ 10)	Log(Sentence)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Above 280g	-0.432	-0.154			0.111	0.459	-0.00540	-0.0250
	(0.297)	(0.670)			(0.302)	(0.452)	(0.0464)	(0.0694)
Above 280g x After 2010	2.359***	2.654**			1.690***	2.234***	0.190**	0.317***
	(0.524)	(1.113)			(0.540)	(0.825)	(0.0908)	(0.119)
Above 50g			0.637***	0.820***	1.369***	1.804***		
			(0.130)	(0.157)	(0.182)	(0.231)		
Above 50g x After 2010			-1.242***	-0.881**	-1.174***	-1.649***		
			(0.275)	(0.362)	(0.435)	(0.458)		
Constant	12.46***	11.13***	9.485***	9.426***	12.02***	15.06***	0.575***	2.225***
	(0.155)	(0.501)	(0.117)	(0.119)	(3.408)	(3.798)	(0.0376)	(0.0524)
Bandwidth	±250g	±50g	±250g	±50g	±250g	±250g	±50g	±50g
Includes Life & <1 Month	No	No	No	No	No	Yes	Yes	Yes
Observations	26,367	2,464	43,642	33,999	26,367	27,478	2,592	2,498
R-squared	0.041	0.017	0.072	0.053	0.045	0.032	0.018	0.020

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. The coefficients in columns 1-2 are estimated from the following regression discontinuity style model:

$$Sentence_{it} = \alpha + \beta_1 Above280_{it} + \beta_2 Amount_{it} + \beta_3 (Above280 \times Amount)_{it} + \delta_1 (Above280 \times After2010)_{it}$$
$$+ \delta_2 (Amount \times After2010)_{it} + \delta_3 (Above280 \times Amount \times After2010)_{it} + g(t)_t + \epsilon_{it}$$

where $Amount_{it}$, the running variable, is the amount of drugs centered at the 280g mandatory minimum, $After2010_{it}$ is a dummy variable equal to one if the case is sentenced after 2010, and $Above280_{it}$ is a dummy variable equal to one if the case involves 280g or more of crack-cocaine. Columns 3-4 estimate equation (1) around the 50g threshold instead of the 280g threshold. Columns 5-6 estimate the sentencing penalty around the 50g threshold and the 280g threshold simultaneously. In addition, all specifications above include a time trend to capture the gradual decline in sentences over time. Column 6 includes life sentences (coded as 70 years) and sentences less than 1 month (coded as 0 years). I typically exclude these sentence types because it is not clear how sentence length should be coded in these cases, but Column 6 show the exclusion of these cases does not affect the results. Column 7 uses a dependent variable that is equal to one if the sentence is greater than or equal to ten years and equal to zero otherwise. Column 8 uses a dependent variable that is equal to the natural log of the offender's sentence. Column 7 includes life sentences and sentences less than one month. In Column 8, sentences less than one month are excluded by the log-operator. I do not find significant differences in these sentencing discontinuities by race. I include the R-squared in this table because the dependent variable is continuous. Figures C1a-g show graphical evidence of the sentencing penalty.

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

Table C2. Sentencing Effects by Year, Reconciling with Bjerk (2017)

	Coefficient on "Just Eligible"		Offense Level (just inelig. / just eligible)	Min{GL,MM} (just inelig./just eligible)
USSC Guidelines Year...	(1)	(2)	(3)	(4)
Years = 2009-2010	0.00328	-1.32	32/32	120/120
	(0.882)	(1.59)		
Year = 2011	0.561	0.916	30/32	97/120
	(0.858)	(1.04)		
Year = 2012	0.657	2.94	30/32	97/120
	(1.46)	(2.00)		
Year = 2013	1.60	1.49	30/32	97/120
	(2.04)	(2.33)		
Year = 2014	4.85***	5.01***	28/30	78/120
	(1.17)	(1.73)		
Year = 2015	10.3***	13.5***	28/30	78/120
	(2.36)	(3.63)		
Years = 2011-2013	0.690	1.59*	-	
	(0.749)	(0.939)		
Years = 2011-2014	1.68**	2.42***	-	
	(0.658)	(0.838)		
Includes Life and < 1 Month	No	Yes	-	

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about data construction. Bjerk (2017) estimates the difference in sentencing between crack-cocaine offenses just above the 280g threshold and offenses just below the 280g threshold using USSC data from fiscal years 2011-2012. He finds no sentencing penalty associated with crossing the mandatory minimum threshold. While there are minor differences in how we construct our analysis sample, I can replicate the results in Bjerk (2017) by limiting the data to fiscal years 2011 and 2012. This is largely because nearly all cases from fiscal year 2011 are sentenced based on the USSC 2010 Sentencing Guidelines. Sentencing guidelines provide a recommended sentence range based on certain facts from the case, such as drug quantity, and prior work has shown that judges follow them closely albeit not completely (Yang 2014). The 2010 Guidelines do not reflect the changes made in the Fair Sentencing Act. This means that for cases from fiscal year 2011 the recommended sentence is the same for cases below versus above the 280g threshold. The Guidelines were updated in 2011 to recommend different sentences for cases above versus below the 280g threshold, reflecting changes made by the Fair Sentencing Act. The Guidelines were updated again in 2014, decreasing the sentence recommendation even further for cases both above and below the mandatory minimum threshold and thus, making the mandatory minimum more binding in cases above the threshold. I estimate the model from Bjerk (2017): $Sentence_{it} = JustEligible_{it} + \epsilon_{it}$ split by USSC Guidelines year (instead of fiscal year). Following Bjerk (2017), **JustEligible** is set equal to one if the weight is from 280-308g and equal to zero if the weight is from 252-280g. I find no sentencing penalty for being above the threshold in cases with the 2009-2010 Guidelines applied (see rows 1-2), a small penalty for cases with the 2011-2013 Guidelines applied (see rows 3-8 and rows 13-14), and an even larger penalty for cases with the 2014-2015 Guidelines applied (see rows 9-12). Column 1 excludes life sentences and sentences less than one month long and column 2 includes these sentences. Column 3 indicates the base offense level for the just ineligible versus just eligible cases. Column 4 indicates the minimum sentence (in months) of the Guidelines recommendation and the mandatory minimum (in months) for just ineligible versus just eligible cases. The sentencing penalty is larger when the difference between the minimum sentences for just ineligible and just eligible cases is larger. This highlights an important point from Bjerk (2017) that the 'bite' of the mandatory minimum is related to the overall federal sentencing structure that crack-cocaine offenders face.
*** p<0.01, ** p<0.05, * p<0.1

Appendix D. Supplementary Materials for Model of Prosecutor Objectives

I. Model of Prosecutor Objectives

First, I detail the prosecutor's decision problem, which determines the probability $Pr(a = mm|s, r)^t$ that a case with a given amount seized s and defendant race r is charged with an amount a that is equal to the mandatory minimum threshold $mm = \{5, 28, 50, 280\}$. Although I do not estimate any of the parameters in the following model directly, I use it to illustrate channels through which $Pr(a = mm|s, r)^t$ may differ by race and to discuss suggestive empirical tests of those various channels.[8]

The prosecutor for the case chooses the **amount (in grams) of drugs charged** a, and can charge amounts higher than **seized evidence** s by collecting additional evidence $a - s$. Seized evidence s is a noisy measure of **true drug trafficking** d, which is unobservable to the prosecutor. For a given case, prosecutor i chooses the amount of drugs charged a to solve the following problem:

$$\max_a \pi(l(a)^t) - \gamma(r, x) \times c_g(a - s) - c_d(|l(a)^t - (l^*(s, r, x) + \phi_i(r, x))|) \tag{2}$$

The function $\pi(.)$ represents the **career benefits** a prosecutor gets from securing a longer sentence. There are also costs to the prosecutor associated with increasing a, such as the **cost of gathering the additional evidence** $c_g(a-s)$ to build the case. This cost $c_g(a-s)$ is increasing in $a-s$.[9] This cost is determined by other actors the prosecutor must face in the process of working a case. Judges, defense attorneys, juries, witnesses, or other actors in the criminal justice system who are racially biased may present fewer obstacles to entering the additional evidence $a-s$ for cases involving black and Hispanic defendants. Also, if defendants of one race procure better defense counsel, that counsel may make it more difficult for the prosecutor to use additional evidence $a-s$. These **cost differences by race** (and other defendant characteristics) are captured in $\gamma(.)$.

The prosecutor also faces a **psychic cost of deviating** $c_d(.)$ from **the sentence that would be justified by law if true drug trafficking were observed** $l^*(d)$. Since true drug trafficking d is unobservable, prosecutors form an expectation of d by solving a signal extraction problem given the seized evidence s, defendant race r, and other characteristics x. This yields $l^*(s, r, x)$.

Finally, a **prosecutor specific taste parameter** $\phi_i(r, x)$ is added to the sentence $l^*(s, r, x)$, reflecting the prosecutor's animus for defendants based on race r or other characteristics x. I assume that only ϕ_i varies at the prosecutor level.

Writing down the prosecutor's objective function makes explicit the various channels that could cause a conditional racial disparity in the probability a defendant is bunched at 280g. First, the disparity could be due to taste-based racial discrimination: $\phi_i(bh, x) > \phi_i(w, x)$. Second, it could be due to statistical discrimination: $l^*(s, bh, x) > l^*(s, w, x)$. Third, it could be due to racial differences in the cost (to the prosecutor) of building a case: $\gamma(bh, x) < \gamma(w, x)$. All three of the channels could also be related to other characteristics x that are correlated with race r rather than race itself.

II. Prosecutor Signal Extraction Problem

The racial disparity in bunching at 280g after 2010 could be due to statistical discrimination. As noted in the main text, statistical discrimination is not innocuous; it is still a form a discrimination and it is not legally permissible. In this section, I outline how statistical discrimination might produce a racial disparity in bunching at 280g in context of the model above. Recall that seized evidence s is a noisy measure of true drug trafficking d. Suppose that, on average, black and Hispanic defendants have higher true drug trafficking amounts: $d_r \sim N(\bar{d}_r, \sigma_d^2)$ and $\bar{d}_{bh} > \bar{d}_w$. Since s is a noisy measure of true drug trafficking d, we can write s as follows: $s = d + \nu,\ \nu \sim N(\mu, \sigma_\nu^2)$

This implies that $E(d|s, r, x) = \bar{d}_r \times (1 - \alpha) + (s - \mu) \times \alpha$ where $\alpha = {}^{\sigma_\nu^2}/_{(\sigma_\nu^2 + \sigma_d^2)}$. Since $\bar{d}_{bh} > \bar{d}_w$, $E(d|s, bh, x) > E(d|s, w, x)$. Since the prosecutor does not observe d, they instead use $l^*(E(d|s, r, x))$. I denote this as $l^*(s, r, x)$,

[8] Note, I write down a static model below, but it can incorporate reputational benefits or reputational costs associated with bunching. The data are not amenable to testing dynamics at the prosecutor-level. I focus on the static problem because it has clear connections to empirical tests I can conduct.

[9] I assume that prosecutors don't suppress evidence or charge amounts lower than the amount seized and thus, $a \geq s$.

and the setting described here implies that $l^*(s, bh, x) > l^*(s, w, x)$. In other words, the prosecutor's expectation over true drug trafficking d "justifies" a higher sentence for black and Hispanic offenders. This decreases their cost of choosing $a > s$ because the associated mandatory minimum sentence will be less of a deviation from that sentence l*. Prosecutors may also use another defendant characteristic x_1 to solve the signal extraction problem (as detailed above) and arrive at $l^*(s, r, x = x_1) > l^*(s, r, x \neq x_1)$. This model offers one potential avenue for statistical discrimination–prosecutor expectations over the true drug amount. Another possible avenue is prosecutor expectations about recidivism risk or the crime-reducing impact of incarceration. I do not observe recidivism outcomes in the data.

III. Prosecutor Responses to Changing Mandatory Minimum Thresholds

The model above also has implications about how the optimal choice in period $t = 0$ (pre-2010) relates to the optimal choice in period $t = 1$ (post-2010). I outline this in Section II.B.2, and provide additional detail in this Appendix section.

Assuming that there are no fixed costs to building a case and that there are no changes in the objective function other than the change in the sentencing schedule, then a prosecutor who chooses not to bunch a case at a mandatory minimum threshold for a sentence X in one period would not bunch the same case at a higher mandatory minimum threshold for a sentence $Y \leq X$ in another period. In other words, a prosecutor not taking on the costs of bunching for a given gain would not take on even greater costs for the same or lesser gain.

For example, when a prosecutor chooses $a^{0*} = s < 5$, this implies that their utility from choosing s is higher than their utility of choosing 5g or 50g: $u(s)^0 > u(5)^0$ and $u(s)^0 > u(50)^0$. Since $a^{1*} = 28$ yields the same benefits as $a^{0*} = 5$ but requires greater costs, then $u(5)^0 > u(28)^0$. These two statements (and the assumptions above) imply that $u(s)^1 > u(28)^1$, which means that the prosecutor should also choose $a^{1*} = s < 5$. The same revealed preference argument can be made for why $u(s)^1 > u(280)^1$. Table D1 shows these possible rational choices of a^{1*} for a given a^{0*} and ranges of s.

Table D1. Relationship between a^{0*} and a^{1*} for relevant ranges of seized evidence

	(1)	(2)	(3)	(4)	(5)
	$s < 5$	$28 > s \geq 5$	$50 > s \geq 28$	$280 > s \geq 50$	$s \geq 280$
$a^{0*} = s$	$a^{1*} = s$	$a^{1*} = \{s, 28\}$	$a^{1*} = s$	$a^{1*} = \{s, 280\}$	$a^{1*} = s$
$a^{0*} = 5$	$a^{1*} = \{s, 28\}$	–	–	–	–
$a^{0*} = 50$	$a^{1*} = \{s, 28, 280\}$	$a^{1*} = \{s, 28, 280\}$	$a^{1*} = \{s, 280\}$	–	–

Ultimately, this means that there will be an increase in the share of cases with $a < 5$ post-2010 (increases from cases previously bunched at 5g and 50g); an ambiguous change in the share of cases with $28 > a \geq 5$ (increases from cases previously bunched at 50g and decreases from cases previously bunched at 5g), an increase in the share of cases with $50 > a \geq 28$ (increases from cases previously bunched at 5g and cases previously bunched at 50g); a decrease in the share of cases with $280 > a \geq 50$ (decreases from cases previously bunched at 50g and cases previously left with $a = s \geq 50$), and an increase in the share of cases with $a \geq 280$ (increases from cases previously bunched at 50g and cases previously left with $a = s \geq 50$). See Figure A2 for a graphical representation of this.

Appendix E. Data Appendix

I. United States Sentencing Commission (USSC) Federal Sentencing Data

These data contain the universe of federal sentences from 1999-2015. The data were obtained from the ICPSR "Monitoring of Federal Criminal Sentences" series here: https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/83.

The data itself is compiled from several court documents: (1) the Judgment and Conviction Order (JC), (2) the Pre-sentence Report (PSR), and (3) the Statement of Reasons (SOR). The PSR is prepared by the probation officer in consultation with the prosecutor and the defense. It is a detailed report on the offender and their offenses intended to aid the judge in making the factual determinations that affect sentencing. The SOR is a form filled out by the judge that details their findings and whether/why they differ from the PSR. The JC is the final ruling in the case that outlines the adjudication and the sentence.

I make several restrictions to these data before using them in analyses. First, I drop any case that is a duplicate in terms of USSC ID number and fiscal year–this only drops one observation. Second, for the main analyses, I remove any case that does not involve a crack-cocaine offense. Third, I remove offenders with a race coded as "other"–this removes people who are recorded as "American Indian\Alaskan Native" (n=346), "Asian or Pacific Islander" (n=387), "Multi-racial" (n=91), and "Other" (n=58). Note, there are people in those race categories who are included because they report Hispanic ethnicity and thus are included as "Hispanic" in the analyses. Fourth, I remove 205 cases that are based on Guidelines amendment years before 1999. Fifth, I remove any case that has a non-zero value for the variable DRUGPROB (1.5% of cases). The USSC codebook explains that this indicates when, "an unresolved problem between the drug type, drug held accountable amount, and base offense level occurs" and that "analyses involving drug weight and base offense level may choose to use this field to remove problematic cases."

Sixth, I remove cases with recorded drug weight above 1000g (about 10% of cases with a weight) to remove outlier cases and cases that involve a dramatically different scale of trafficking. This also removes cases where drug weight is coded as a range (about 19% of cases) or where drug weight is missing (about 2.6% of cases). I explore robustness to these restrictions in Appendix A. Seventh, I remove cases that have missing values for any of the key variables used in the main analyses: district, male, college, criminal history, age, citizenship, number of current offense counts, sentence length, number of dependents, race, and year. This removes about 2% of cases. Eighth, I remove extreme outliers in sentence length by calculating the difference between the Guidelines recommendation and the actual sentence and removing cases where this difference is in the top 0.5% of observations. Ninth, I remove cases where the judge does not accept the facts of the case. From 2006-2015, I make this restriction using the CHP2CHG variables, which allow me to exclude cases based on whether the judge specifically took issue with the drug amount. From 1999-2005, that variable is not available, so I use the ACCGDLN variable, which allows me to exclude cases based on whether the judge took issue with any Guidelines factor. This removes about 6% of cases. Note, the main result is robust to ignoring this restriction, judge acceptance of facts does not change sharply after 2010, and there is no racial difference in whether a judge accepts the facts of the case before or after 2010. Finally, I remove cases where the SOURCES variable does not equal one. The USSC describes these cases (SOURCES=1) as indicating that the "information represents known court findings." The USSC notes, "when conducting guideline-based analysis, it is preferable to include cases where guideline application is clearly documented by the court. Including only [SOURCES] = 1 cases in the analysis will minimize, but not eliminate, data inconsistencies." This removes about 4% of cases. This results in a final sample of 50,273 cases.

Again, the main results are robust to all of these restrictions and factors that I restrict on do not change sharply in 2010. Around 29.3% of the sample is dropped in 2010 and 29.6% is dropped in 2011. Around 34.3% is dropped in 2009 and 35.3% is dropped in 2012. Beginning with the full sample of crack-cocaine cases and excluding only those cases with a missing weight or weight recorded as a range yields the following racial disparity in bunching at 280-290g after 2010: β=.022, SE=0.005, n=66,357. This sample is 79% of the full sample of crack-cocaine cases. Including cases coded as a range and using the lower bound of the range as the weight yields the following racial disparity in bunching at 280-290g after 2010: β=.026, SE=0.008, n=81,944. This sample is 97% of the full sample of crack-cocaine cases. Including cases coded as a range (coded as described above) and cases with a missing weight but inferring a lower bound of the weight from the Base Offense Level yields the following racial disparity in bunching at 280-290g after 2010: β=.024, SE=0.008, n=84,094. This sample is 99.8% of the full sample of crack-cocaine

cases. The remaining cases are not in the analysis because they are missing information about offender race. Coding those missing values as white yields the following racial disparity: β=.025, SE=0.007, n=84,230. Coding those missing values as black or Hispanic yields the following racial disparity: β=.024, SE=0.008, n=84,230. With those missing values recoded the sample is 100% of the full sample of crack-cocaine cases from fiscal years 1999-2015. Ultimately, the estimated disparity is similar with the restrictions outlined above and without them. Key variables from the data are described below:

Crack cocaine offense. Whether or not the case involves a crack cocaine offense is derived from the raw variables DRUGTYP{X} provided by USSC. These variables contain the types of drugs involved in the offense. This information is taken from the Judgment and Conviction Order (JC), if present. If it is not included in the JC, the information is taken from the Pre-sentencing Report (PSR) prepared by the probation officer assigned to the case.

Drug quantity. The amount of drugs involved in the case is derived from the variables WGT{X} provided by the USSC. These variables contain the gram amount for drug {X} corresponding to DRUGTYP{X}. I use the weight corresponding to the drug type crack cocaine for each case. Information on drug amount is taken from the Statement of Reasons (SOR), if present. If not present in the SOR, the information is taken from the PSR.

Offender race. I code offender race based on the USSC variables NEWRACE, which categorizes offenders as non-Hispanic white, non-Hispanic black, or Hispanic. The variable NEWRACE is a combination of raw variables MONRACE and HISPORIG. The information for these variables is taken from the PSR. The USSC notes that offender race is self-reported to the probation officer.

Other offender characteristics (e.g. education). These are also derived primarily from the PSR.

Year. The year used for analyses is derived from the variable AMENDYR, which represents the year of the USSC Guidelines manual used for sentencing guidelines calculations. This information is taken from the PSR.

District. The district used for analyses is derived from the variable DISTRICT, which represents the federal district the offender is sentenced in. This information is taken from the JC, if available, and from the PSR, if not.

II. FL State Inmate Database

These data contain all inmates who have been released from a FL state prison since October 1997. The data were obtained here: http://www.dc.state.fl.us/pub/obis_request.html.

I downloaded this data in April 2016; the current database available will differ somewhat. In this data, I remove duplicates that are created by appending data on active and released inmates and remove around 4,000 records where the charge is recorded as "adjudication withheld." Finally, I only keep offenses involving cocaine. Key variables from the data are:

Offense/drug quantity. The offense field indicates all of the inmate's known offenses in FL. For drug offenses, the field contains the drug name. In FL, powder-cocaine and crack-cocaine cases are both recorded as "cocaine." For many of the drug offenses, the field contains a label indicating if the offense was with 0-28g of cocaine, 28-200g, 200-400g, or 400+g.

Offender race. Offender race is included as part of the "basic inmate information" file. There is no information on how race is determined. I expect it is similar to the federal court data, in which race is self-reported. In the FL data, the race field includes labels for "black", "Hispanic", and "white" inmates.

These data cover cases that are handled at the state/local level as opposed to federal court (those cases included in the USSC data). This is important because state and local authorities could send more of their high weight, 280g cases to federal court after 2010. Similarly, federal prosecutors could pull more of these types of cases from state and local courts after 2010. A case can enter the federal system for procedural reasons: drugs are trafficked across state lines or the arrest is made by federal agents. However, cases can also be prosecuted federally for more arbitrary reasons. Wright (2006) notes that sorting into federal versus state is often determined by law enforcement agents involved with the case and/or the prosecuting attorneys, but it is never the official purview of judges or defense attorneys. Why might local law enforcement or attorneys wish to pass a case on to the federal courts? For one, local authorities may not have the time or resources to properly pursue a case. Also, Wright suggests that federal sentencing is typically harsher than state sentencing, and that this gap could motivate jurisdiction decisions.

III. NIBRS Property Segment

These data contain information on drug quantity and drug type for drugs seized by NIBRS-participating police departments. The data were obtained here: https://www.icpsr.umich.edu/icpsrweb/NACJD/series/128.

I make a few restrictions on this data before using it in analyses. For tractability, I limit the offender segment to incidents that involve 5 or fewer offenders. This covers 99% of all incidents. Also, the fraction of incidents with 5 or fewer offenders does not meaningfully change after 2010 (99.1% in 2000, 99.1% in 2005, 99.0% in 2010, and 99.3% in 2015). Finally, it is not correlated ($\rho = 0.0001$) with the probability an incident involves 280-290g of crack-cocaine. Next, I remove observations that are missing race. I then collapse the data to the incident level, summing total quantity seized to the incident level. I also remove cases missing other key variables: sex, age, and state. For the primary analyses of the NIBRS data, I limit the sample to a balanced panel of agencies. For robustness checks, I limit to states that have had full agency coverage in NIBRS since 2012 and over 90% coverage since 2008. Finally, I only keep cases with weights below 1000g. Key variables from the data are outlined below:

Drug quantity. The drug quantity field is populated when there is a drug seizure by the department. It is equal to the total quantity of drugs seized.

Offender race. The race field for NIBRS does not include an indicator for whether the offender is Hispanic. An ethnicity field is available only in later years, so I focus on white versus black offenders in this data. There is no information on how race of the offender is determined. I expect it is similar to other criminal justice data (e.g. the USSC data), in which race is self-reported.

IV. DEA STRIDE Database

These data contain information on drug quantity, drug type, and purity for seizures and undercover purchases sent to DEA labs for analysis. The data also indicate whether the drugs were obtained via seizure or undercover purchase. For drugs that were purchased, the data contains their price. The data were obtained from a FOIA request for all records related to cocaine from January 1, 1999 to December 31, 2015. I only keep cases with weights below 1000g. Key variables from the data are:

Drug quantity. This field indicates the weight of the drug evidence received by the lab.

Drug type. This field indicates type of drug. The DEA does not use street names to refer to drugs in this data, meaning no drugs are referred to as crack-cocaine. For the main analyses, I use all drug types containing the word "cocaine," but results are similar if I focus on the "cocaine base" drug type.

Purity. This field indicates the chemical purity of the drug evidence received by the lab.

Acquisition. This field indicates whether the drug was acquired via seizure or undercover purchase.

Price. This field is populated if the drugs were acquired via undercover purchase. Price indicates the price paid for the drugs. In one robustness analysis, I plot the time series of price by month. To do this, I adjust the raw price field (described here) based on the purity of the drug, calculating a "price per pure gram."

V. EOUSA Case Management Files

A. Data Description

These data contain information on cases handled by the EOUSA from the EOUSA's internal case management system: Legal Information Office Network System (LIONS). The data were obtained here:
https://www.justice.gov/usao/resources/foia-library/national-caseload-data.

For the main analyses, I limit to cases with weights below 1000g. This removes a large portion of the sample for which drug weight is missing. I discuss missing values in Section B below. Key variables from the data are:

Drug quantity/type. This field comes from the "controlled substances" screen of the LIONS software. According to the LIONS user manual (EOUSA 2016), the controlled substances data "tracks information on controlled substances; includes type and quantity of all substances in a case." The manual instructs users to do the following: "Enter the actual quantity of the controlled substance seized. Fractions must be converted to one or two decimal places." The software itself, however, simply has a field for "quantity" to be entered with no instruction. In general, the drug weights recorded in the EOUSA data are much larger than the drug seizure weights reported by the DEA or

NIBRS. In fact, drug quantities decrease in the DEA and NIBRS after 2010 but increase in the EOUSA. Also, the fraction of 280-290g cases at the district/month level in the EOUSA data is correlated with the fraction of 280-290g cases at the district/month level in the USSC data. These validation tests suggest the data entered into LIONS is indicative of total drugs involved/charged in the offense and not raw amount seized alone. Rehavi and Starr (2013) note an issue with the EOUSA data on drug quantities beginning in 2004, however, it appears that issue may be corrected in later versions of the EOUSA data. Looking at districts with non-missing quantity information from 2000-2015, the correlation between mean quantities in the USSC data and the EOUSA data is the same pre- vs. post-2004 and a line plot suggests the yearly means track each other well over time (see Figure E1a). Finally, density plots of the drug quantities from EOUSA for crack-cocaine from 1999-2003 versus 2005-2009 are very similar (see Figure E1b and E1c).

Another important question about this field is: when is this information entered into the system? The LIONS User Manual instructs attorneys to initiate a criminal file when: (1) "an investigation is authorized and an attorney spends one hour or more working on the investigation"; (2) "a citation, complaint, information or indictment is filed in court"; or (3) "a case on appeal, in which the United States is a defendant, is remanded to the District Court for a new trial." This indicates that for new cases, the initial information is entered shortly after the case is received or after the first court filing is made. In other words, the initiation of the file should begin long before final sentencing occurs. When a case is initiated, the user is guided through a number of screens and instructed to enter information. One of the screens that the user encounters upon initiating a criminal file is the "controlled substances" screen with the fields: type, quantity, measure, and other. These are not required fields and can be left blank. In a small fraction of cases, there are multiple weights recorded for crack-cocaine. In the main analyses, I use the minimum of these weights. The main results (i.e. the bunching magnitude, the *Alleyne* effect, and the missing mass estimates) are all robust to using the sum of those weights or the maximum of those weights.

The timing of file initiation outlined in the LIONS User Manual and the presence of the "controlled substance" screen upon initiation, indicates that this information is first entered shortly after a case is received or the first court filing is made. This information can be updated, however, this is not required. There are many events that occur during the life of a case that the attorney may wish to record in LIONS. For example, if the case is assigned to a new attorney or a superseding indictment is filed, it may be important to enter this information into LIONS for management or statistical reporting purposes. Note, the User Manual never indicates that the controlled substances information is used for broader management purposes, like case assignment, and the statistical reports released by the EOUSA do not include controlled substances information. Users are also instructed to enter new information when an appeal begins. At no point in entering new information is the user prompted to revisit and update the controlled substances information. The controlled substances screen is only accessed when first initiating a case or by directly choosing to update the controlled substances information.[10]

Staff ID/Assignment. The EOUSA data also contains an ID for the lead attorney assigned to the case. This ID is tied to the district. Two attorneys can have the same ID as long as they are in different districts. Staff ID is always present for the lead attorney on the case. The LIONS User Manual notes, "At a minimum, you must track the Lead Attorney. This assignment information is used by LIONS to create individual and master calendars."

Initials. Since the EOUSA staff ID for lead attorney is not constant across districts, I use a field for the attorney's "initials" to follow attorneys who switch districts. The initials field is "initials of the staff member authorized to use the LIONS application." In most cases, the field contains 3 or more letters, making it likely that if I see the same initial in two different districts it is the same attorney. This initials ID appears to accurately identify attorneys who switch districts. First, attorneys who move from one district to another continue to bunch in the new district. Second, when an attorney moves into a new district, other attorneys in that district start to bunch more. Third, attorneys who I identify as "moved" are often disconnected from their old district in the data and connected to their new district. Fourth, the Figure A16 results are robust to excluding attorneys with initials that only contain 2 characters, for whom I am more likely to identify a false positive move. Fifth, the Figure A16 results are robust to excluding attorneys who have a time between districts that is greater than one year, which could also be indicative of a false positive. These results are available upon request. If the initials ID were random, we should not expect to see these patterns.

[10] Such updates can occur; approximately 3% of cases received in the year before the FSA but sentenced after the FSA end up with a quantity bunched at 280g. However, these updates are not the norm–for comparison, in cases both received and sentenced after the FSA, 20% have a quantity bunched at 280g.

Date received. The date the criminal case was received by the US Attorney's Office.

Sentence date. For cases that are sentenced, the EOUSA also notes the data of sentencing.

Judge ID. For cases that are brought to a judge, the EOUSA data contains an identifier for the judge involved and that identifier can be linked to a table of judge names. For robustness analyses, I examine the effect of judge race and political party on bunching at 280g. I obtain data on judge characteristics from Crystal Yang's paper on resource constraints and judicial vacancies (see Appendix References for link).

B. Missing Values

Many fields in the EOUSA data have missing values. Although some fields are required (e.g. case ID, lead attorney ID, instrument type), the LIONS case management software does not require the user to enter information in all fields. The drug quantity field on the controlled substances screen is not required. Approximately 60% of crack-cocaine cases are missing quantity information. This varies across districts, with some districts missing quantity in as little as 7% of cases and some missing it in as many as 99.8% of cases. Importantly, all of the main results from the EOUSA data are robust to the inclusion of district fixed effects. The fraction missing does increase after 2010 in the EOUSA data. Specifically, it increases from 60.7% in 2010 to 66.2% in 2011. Because of this sharp increase in missing values that coincides with timing of the FSA, I explore robustness to various ways of dealing with missing values in Appendix A and in this supplementary Appendix.

In Appendix A (and discussed in the main text), I make the assumption that all cases with a missing drug weight would be recorded outside of the 280-290g range. This assumption produces a time series pattern of bunching that is similar in dynamics and magnitudes to the pattern of bunching found in the USSC data. In Table A15 and Figure A12b, I show that the increase in bunching after 2010 is robust to this imputation of missing values. Figure A13 shows how calculating the fraction of bunching attorneys in the data is affected by this imputation of missing values. Table A18 shows that bunching across districts is persistent–attorneys that move from one district to another tend to bunch at the mandatory minimum threshold in both districts. Since this involves cross-district comparisons and missingness varies by district, Table A18 shows that this result is also robust to imputing missing values as zero. The main results on the effect of moving attorneys on the receiving districts in Figure A16 also impute missing values as zero; again, this is because that analysis involves cross-district comparisons.[11] Finally, Table A24 shows that the prevalence of missing values does not increase after *Alleyne*, and that the decrease in bunching after *Alleyne* is robust to imputing missing values as zero. In general, the main results are all robust to assuming cases with missing values are cases that would not be recorded in the 280-290g range.

In this section, I consider additional analyses using a few alternative approaches to dealing with missing values. First, in Figure E3, I introduce five alternative methods for dealing with missing values in the EOUSA data. In panel (a), I keep the missing values as-is but exclude districts that have missing values in more than 75% of cases. With this restriction, only 40% of total cases are missing quantity. In panel (b), I keep the missing values as-is but exclude districts that have missing values in more than 50% of cases. With this restriction, only 30% of total cases are missing quantity. In panel (c), I use the counts of 280-290g cases in the USSC data (excluding cases coded as a range) pre- and post-2010 to determine how many missing value cases I should randomly impute as In280290=1.[12] This assumes that the 280-290g cases are missing at random. Panel (d) takes a similar approach as panel (c) but includes cases coded as a range in the USSC counts. Finally, in panel (e), I assume that missing 280-290g cases (based on the USSC counts excluding cases coded as a range) are not assigned to any attorneys who I otherwise classify as bunching attorneys (based on their cases with non-missing values). This is an extreme assumption. In practice, bunching and non-bunching attorneys have the same fraction of cases with missing values. Also, comparing EOUSA counts of 280-290g cases to USSC counts of 280-290g cases indicates that 280-290g are less likely to be missing than cases with other amounts. Comparing counts between the datasets suggest the EOUSA is only missing quantity information for 47% of 280-290g cases whereas it is missing quantity for 60% of cases overall. Nevertheless, all five panels indicate an increase in bunching at 280-290g after 2010 and, with the exception of panel (d), a decrease in bunching at 280-290g after 2013.

[11] The results are robust to excluding missing values. See Figure E3.

[12] Assigning missing cases to 280-290g at random means that the results may differ depending on the random draw. For inspection, I run 10 replications and find that the results do not vary markedly across replications. I report the result of the first replication.

Another main result from the EOUSA data is in classifying the fraction of attorneys who have cases bunched at 280-290g. Dealing with missing values changes this fraction somewhat, but not dramatically. Assuming that 280-290g cases are missing at random or even that they would all be assigned to attorneys not classified as "bunching," leads to a fraction of bunching attorneys around 50%. Limiting to districts with fewer missing values leads to a fraction of bunching attorneys around 17-30%. Finally, another key result from the EOUSA data is the decrease in bunching after *Alleyne*. Table E1 shows that the decrease in bunching after *Alleyne* is robust to these alternative ways of dealing with missing values.

VI. Additional Datasets

National Survey on Drug Use and Health (NSDUH), 2002-2016.

These data contain survey responses to questions about drug use. The data were obtained here: https://www.datafiles.samhsa.gov/study/national-survey-drug-use-and-health-nsduh-2002-2016-nid18454.

Key variables from the data are:

Crack-cocaine use. This variable is based on the question, "have you ever, even once, used crack-cocaine?" It is coded as one if the respondent indicates ever using crack-cocaine and zero otherwise.

Drug selling. This variable is based on the question, "during the past 12 months, how many times have you sold illegal drugs?" It is coded as one if the respondent indicates any drug selling in the past 12 months and zero otherwise. I also use this variable in conjunction with the crack-cocaine use variable to create an indicator for whether the respondent has ever used crack-cocaine **and** has sold drugs in the past 12 months.

Race. This variable represents the respondent's self-reported race as non-Hispanic white, non-Hispanic black, or Hispanic. I remove respondents who report a race that does not fall in these three categories.

Google Search Trends Data on Racial Animus from Stephens-Davidowitz (2014), 2004-2007

These data contain a measure racial animus at the state-level introduced by and described in detail in Stephens-Davidowitz (2014). The data were obtained here: http://sethsd.com/s/StateIATGoogle-4say.csv. Key variables from the data are:

Racial animus, Google. Stephens-Davidowitz uses Google search data from 2004-2007 (accessed via Google Trends) and measures relative search volume in every US state for a specific racial slur and its plural form.

Implicit Association Test (IAT) Data on Racial Animus for Lawyers, 2006-2016.

These data contain the results of implicit association tests for racial bias for over 3 million individuals. Although recent research casts doubt on the validity of the IAT for detecting bias (Oswald et al. 2013), the data has two advantages. First, it can be aggregated to the federal district, a sub-state geography. Second, it can be calculated solely for people reporting an occupation of "Lawyers, Judges, and Related Workers." The data were obtained here: https://osf.io/52qxl/. In the process of merging the county-level data to district-level codes, some state codes are excluded (AA, AE, AP, AS, FM, GU, MH, MP, PR, PW, and VI). Also, some small municipalities in AK are excluded and 41 counties are not represented in the IAT data. Key variables from the data are:

Overall IAT score. This is measured by having a person sort words into "good" and "bad" categories, sort people into "black" and "white" categories, and finally, sort both words and people into "black" and "white" categories paired with "good" or "bad" categories. The time it takes to sort into "black/good" relative to "black/bad" and "white/bad" relative to "white/good" is the basis of a person's score. See "Project Implicit" for more detail.

Occupation. This variable is self-reported in the course of taking the IAT. I use this to limit the scores to people with an occupation of 23-1000, "Legal - Lawyers, Judges, and related workers." Note, this is different from the occupation 23-2000, "Legal - Legal support workers" and 33-3000, "Protective services - Law Enforcement."

County and state. This variable is self-reported in the course of taking the IAT. I use these geographic location variables to aggregate data to the district level.

Survey of Inmates in Federal Correctional Facilities (SIFCF), 2004.

These data contain survey responses from people incarcerated in Federal prisons in 2004. The data were obtained here: https://www.icpsr.umich.edu/web/NACJD/studies/4572. Key variables from the data are:

Involvement in drug organization. Based on the question, "In the year before your arrest on [insert date], were you a part of any group or organization that engaged in drug manufacturing, importing, distribution or selling?"

Illegal income. Based on the question, "During the month before your arrest, [...] Did you receive any income from illegal sources?"

Type of counsel. Based on the question, "Did you hire your own lawyer or was one assigned by the court?"

Drug weight. Based on the question, "Approximately what amount of crack cocaine was involved?"

Number of members in drug organization. Based on the question, "Altogether, about how many people would you say regularly participated in that group or organization?"

Role in drug organization. Based on the question, "Which of these best describes your role in that group or organization– (1) A leader or organizer? (2) A middle man? (3) An underling, such as a carrier, runner, etc? (4) A seller? (5) Other." I use the responses to form binary variables indicating the inmate's role in the group.

Race. Based on self-reported race for non-Hispanic white, non-Hispanic black, and Hispanic inmates. I remove any respondents who report a race outside of one of those three categories.

United States Sentencing Commission (USSC), Appeals Data, 2012-2013.

The USSC provides data on cases appealed in the 12 circuit courts of appeal. I focus on cases with appeal dates in fiscal year 2012 and 2013. The data were obtained from here: https://www.icpsr.umich.edu/web/NACJD/series/75. Key variables from the data are:

Guideline issue appealed. This field indicates the issues that are appealed from the original case. I use this field to determine if any drug issue is appealed, if any issue with the drug amount is appealed, and if any issue with the application of the mandatory minimum is appealed.

Appeal disposition. This field indicates the outcome of the appeal. I use this field to determine if the issue is reversed or remanded to the lower court.

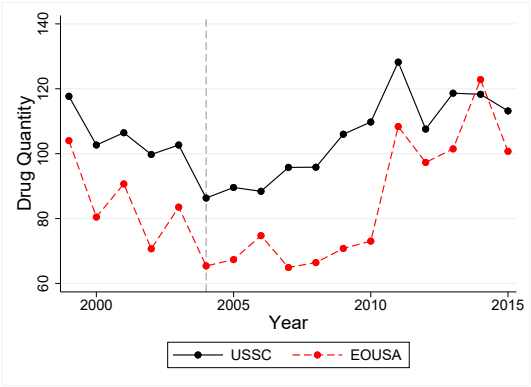
District. This field indicates the district where the case was originally sentenced.

Year. This field indicates the fiscal year in which the appeal is disposed.

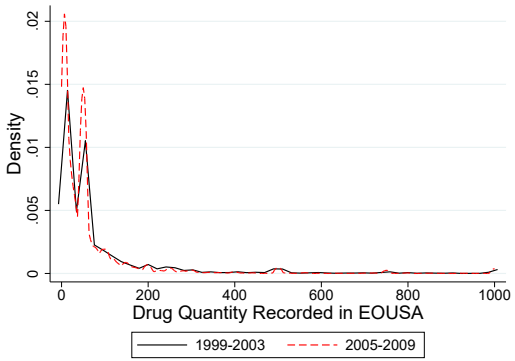
I use the variables above to determine if the district has any drug appeals from 2012-2013, if it has any drug appeals with an issue with the drug amount, if it has any with an issue with the mandatory minimum application, if it has any of those issues + is reversed, and if it has any of those issues + is reversed or remanded to the lower court.

Figure E1. Comparison of Drug Quantity Data in EOUSA, Pre- vs. Post-2004

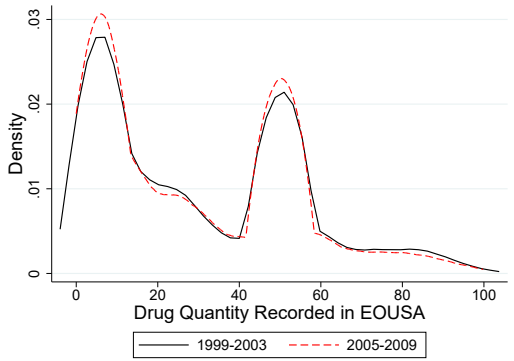
(a) Mean Quantities in EOUSA vs. USSC



(b) Pre-2004 vs. Post-2004 Density, 0-1000g



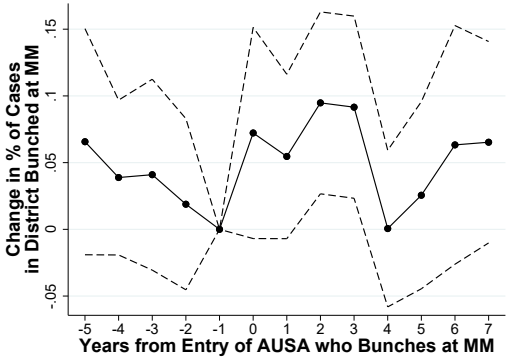
(c) Pre-2004 vs. Post-2004 Density, 0-100g



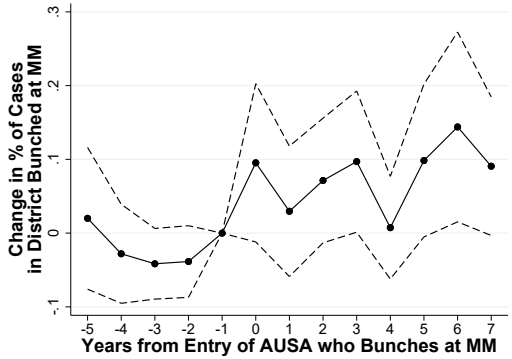
Notes. Figure E1a plots the mean drug quantity for crack-cocaine cases from the EOUSA data in each year and the mean drug quantity for crack-cocaine cases from the USSC in each year. These series track each other well and the correlation does not change in 2004. The only major change is in 2010 when the bunching can be seen more prominently in the mean values from the EOUSA than in the mean values from the USSC. Figure E1b plots the density of drug quantities in crack-cocaine cases from the EOUSA data for years 1999-2003 and years 2005-2009. In this version of the EOUSA data, there is no major change in these distributions post-2004. Figure E1c zooms in on the 0-100g range to show this more clearly.

Figure E2. Additional Evidence of Prosecutorial Discretion in Bunching, Movers Results, EOUSA

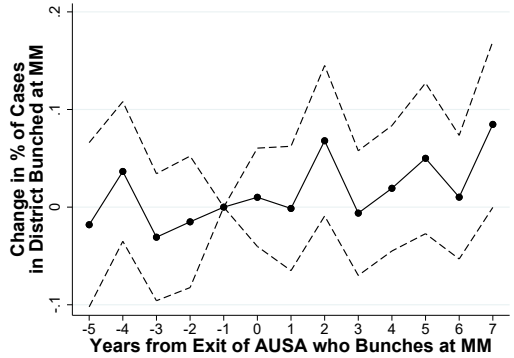
(a) Effect of Entry of a Bunching AUSA



(b) Effect of Entry of a Bunching AUSA, Low-Bunching Districts



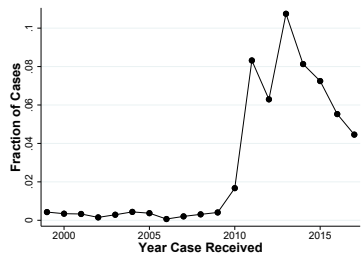
(c) Effect of Exit of a Bunching AUSA



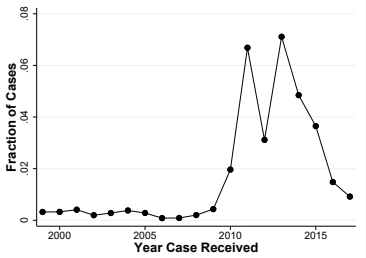
Notes. See Figure A16 notes. Since this analysis involves cross-district analysis and missingness varies across districts, I impute missing weights as zero instead of excluding them from the analysis for Figure A16. These figures exclude the missing weights. The post-period increase in is 0.029, p-value=.18 for all districts (panel a) and is 0.098, p-value=0.002 for low-bunching districts (panel b). These plots are created from the EOUSA data.

Figure E3. Main Bunching Result in EOUSA Data, Various Missing Value Recodings

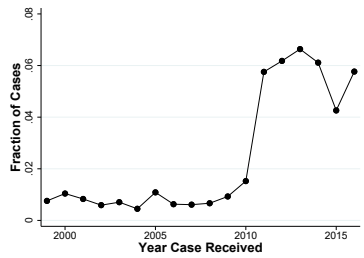
(a) Excludes Districts with >75% Missing



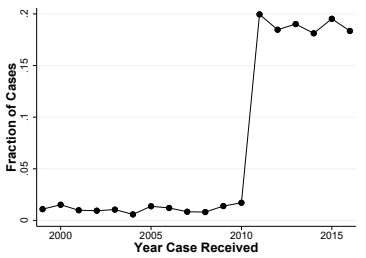
(b) Excludes Districts with >50% Missing



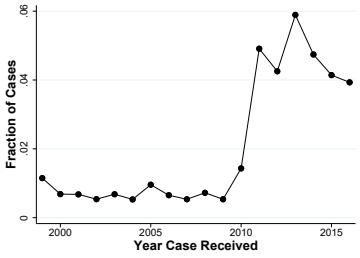
(c) Imputing Missing 280-290g At Random Based on USSC Counts (Excluding Ranges)



(d) Imputing Missing 280-290g At Random Based on USSC Counts (Including Ranges)



(e) Assigning Missing 280-290g to Attorneys Not Classified as Bunching



Notes. These figures plot the fraction of cases recorded with 280-290g in the EOUSA data. Each panel uses a different approach for dealing with missing values in the EOUSA data. Panel (a), keeps the missing values as-is but excludes districts that have missing values in more than 75% of cases. Panel (b) keeps the missing values as-is but excludes districts that have missing values in more than 50% of cases. In panel (c), I use the counts of 280-290g cases in the USSC data (excluding cases coded as a range) pre- and post-2010 to determine how many 280-290g cases are missing in the EOUSA data, and I use that difference to randomly impute cases as In280290=1. This assumes that the 280-290g cases are missing at random. Panel (d) takes a similar approach as panel (c) but includes cases coded as a range in the USSC counts. In panel (e), I assume that missing 280-290g cases (based on the USSC counts excluding cases coded as a range) are not assigned to any attorneys who I otherwise classify as bunching attorneys (based on their cases with non-missing values).

Table E1. Effect of *Alleyne v. US*, Further Missing Robustness

			280-290g		
	(1)	(2)	(3)	(4)	(5)
After June 17th, 2011-2016	0.0279	-0.0009	0.0089	0.0062	-0.0069
	(0.0278)	(0.0244)	(0.0136)	(0.0222)	(0.0124)
After June 17th, 2013	-0.1905**	-0.1951*	-0.0663**	-0.0838*	-0.0376
	(0.0877)	(0.1060)	(0.0290)	(0.0487)	(0.0277)
Constant	0.1866	-0.0034	0.1048**	0.1499***	0.0581*
	(0.1774)	(0.0105)	(0.0448)	(0.0576)	(0.0319)
Specfication	Excluding Districts with > 75% missing	Excluding Districts with > 50% missing	Recoding at random based on USSC counts	Recoding at random based on USSC, including ranges	Assigning all missing to 'non-bunchers' based on USSC
Bandwidth	±150 days	±150 days	±150 days	±150 days	±150 days
Observations	1,667	1,160	6,182	6,182	6,182

Notes. Standard errors clustered at the date the case is received in parentheses. The estimates in this table are based on the EOUSA data. See Table 7 for notes on estimation. Column 1 excludes districts with missing quantity for more than 75% of cases. Column 2 excludes districts with missing quantity for more than 50% of cases. Column 3 imputes missing value cases at random based on counts of 280-290g cases in the USSC data, excluding cases coded as a range. Column 4 imputes missing value cases at random based on counts of 280-290g cases in the USSC data, including cases coded as a range. Column 5 assigns missing 280-290g cases to attorneys not classified as bunching attorneys based on counts of 280-290g cases in the USSC data, excluding cases coded as a range. Note, the estimated decreases in bunching after *Alleyne* in Columns 3-5 are comparable in size to the decrease in Table A24, where I impute all missing value cases as In280290=0.

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

Appendix F. Alternative Methods of Estimating Bunching

I. Comparing Aggregated Pre- and Post-2010 Densities

Most papers using the "difference-in-bunching" approach can be fit into one of two categories. In one, authors estimate bunching using the conventional polynomial method (see section II below for a detailed description) separately for groups where the threshold applies and for groups where the threshold does not apply, using the latter as a placebo test (e.g. Fack and Landais 2016; Gelber, Jones, and Sacks 2020; Zaresani 2020). In the other, authors directly compare the group where the threshold applies to the group where the threshold does not apply. Even within the direct comparison category, strategies differ. Several papers compare the distributions by aggregating the data into bins and calculating the difference in levels between the actual and the counterfactual distributions (e.g. Brown 2013; Cengiz, Dube, Lindner, and Zipperer 2019; Best et al. 2020). Others compare the distributions using regression analysis on the microdata (see citations in main text). These papers frequently estimate the difference in the probability an observation is in a given bin between the actual and the counterfactual setting.

In this paper, I am primarily interested in estimating the change in the probability a case is charged with 280-290g after 2010 and whether that change in probability differs by race. In addition, some analyses in the paper preclude aggregating the data into bins because they rely on data that do not include precise drug quantities. For these reasons, I follow the papers that use regression analysis on microdata to compare the pre- and post-2010 crack-cocaine distributions.

To show robustness to the other "difference-in-bunching"/direct comparison method, I aggregate the cases into 10g bins pre- and post-2010. Following Best et al. (2020), I estimate 90% confidence intervals with a bootstrap procedure that samples cases with replacement from the microdata before aggregating to the 10g bin level.[13] I compare the binned distributions to estimate the net change in bins below 280g, at 280-290g, and above 290g.

Aggregate bunching analyses yield very similar results. Figure F1 below plots the counterfactual scaled pre-2010 density and the actual post-2010 density. The spike at 280g in the post-2010 density is the bunching that is detected in Table 2. After 2010, there is a 3.3 percentage point increase in cases with 280-290g. I also show the densities by race. The bunching at 280g in the post-2010 density is larger for black and Hispanic offenders. After 2010, the rise in cases with 280-290g is about 2 percentage points higher for black and Hispanic offenders than for white offenders. Table F1 displays the results from similar binned analyses using the NIBRS data, DEA data, and EOUSA data.

II. Comparing an Estimated Counterfactual and Post-2010 Densities

Many bunching papers, for lack of variation in the threshold of interest, estimate bunching by constructing the counterfactual density from the actual bunched density. To do this, one typically aggregates the data into bins and estimates a regression of the count in each bin on a high-order polynomial of the bin's value and dummy variables for bins in the bunched "window." The estimates from that regression (not including the bunching dummy variables) can be used to predict a smooth distribution of bin counts. Authors then compare that smooth density to the actual density to calculate the degree of bunching in the actual density. My main results are also robust to this method.

To start, I collapse the data on drug quantities for all cases after 2010 to 10 gram bins. I then run a regression of the count of cases on a seventh order polynomial of the bin values and dummy variables for the bins 0-10g, 270-280g, and 280-290g. Then, using the coefficients from the seventh order polynomial and the dummy variable for the bin 0-10g, I calculate a smooth counterfactual distribution. For graphical purposes, I re-scale that smooth distribution to have the same total number of cases as the true distribution. Next, I calculate the percent of all cases that are in the 280-290g bin in the true distribution, the percent of all cases that are in the 280-290g bin in the counterfactual distribution, and the difference between those two percentages. Finally, I run a regression of the difference between the true and counterfactual distributions on a dummy variable equal to one for the 280-290g bin and equal to zero otherwise (bootstrapped standard errors are calculated by re-sampling the residuals from the polynomial estimation with 200 replications). I carry out a similar procedure to estimate the difference in bunching between white and

[13] I draw 50 random samples from the microdata and do the binned analysis on each sample. The final number of cases for each bin is calculated as the mean of the number of cases across all 50 samples, and the final standard error is calculated as the mean of the standard error across all 50 samples.

black and Hispanic offenders (the major difference being that I estimate the counterfactual distributions separately for white and black and Hispanic offenders and that the final regression includes an interaction between the 280-290g bin dummy and a dummy for black and Hispanic offenders).

First, I construct the counterfactual density by aggregating the data to 10-gram bins, summing the number of cases in each bin. With this aggregated data, I estimate a regression of the bin counts on a seventh-order polynomial of the bin values, dummies for the 270g and 280g bins, and a dummy for the 0g bin.

$$Count_b = \alpha_0 + \sum_{i=1}^{7} \beta_i (Amount_b)^i + \gamma_1 Bin270_b + \gamma_2 Bin280_b + \delta_1 Bin0_b + \epsilon_b \tag{3}$$

where $Count_b$ is the total number of cases in bin b, $Amount_b$ is the value of bin b, and $Bin[X]_b$ is a dummy variable indicating if the bin's value equals X. I use the parameter estimates from (1) (excluding γ_1 and γ_2) to predict a smooth density of bin counts. Furthermore, I adjust the predicted counts to force the smooth density to have the same number of cases as the actual density. I plot the counterfactual density and the actual post-2010 density below in Figure F1.

Using the predicted counts from the counterfactual density and the actual counts post-2010, I construct the percent of cases in each bin for each density. I then calculate the difference in these percentages and run the following regression, bootstrapping the standard errors from 200 replications:

$$(\%\text{ in Post2010 - }\%\text{ in Predicted})_b = \alpha + \beta Bin280_b + \epsilon_b$$

The resulting $\beta = 0.0366$ and $SE_\beta = 0.0178$.

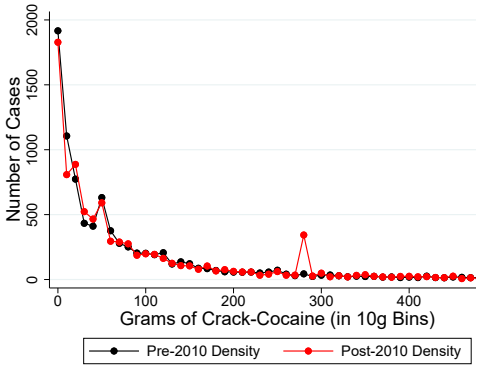
Next, I estimate:

$$(\%\text{ in Post2010 - }\%\text{ in Counterfactual})_{br} = \alpha + \beta Bin280_b + \gamma NonWhite_r + \delta Bin280_b \times NonWhite_r + \epsilon_b$$

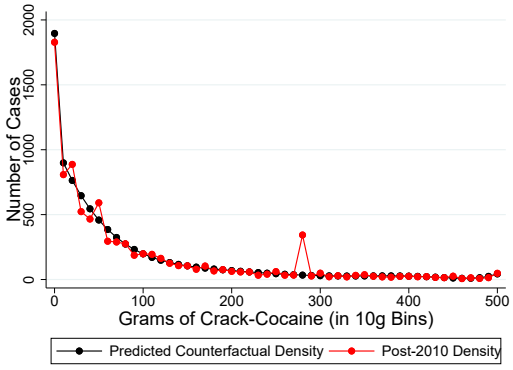
Using the Saez (2010) and Chetty et al. (2011) method, I estimate $\delta = 0.0237$ and $SE_\delta = 0.0117$. In all analyses, I detect substantial bunching after 2010 and disproportionate bunching after 2010 for black and Hispanic offenders.

Figure F1. Counterfactual Distributions vs. Post-2010 Distributions

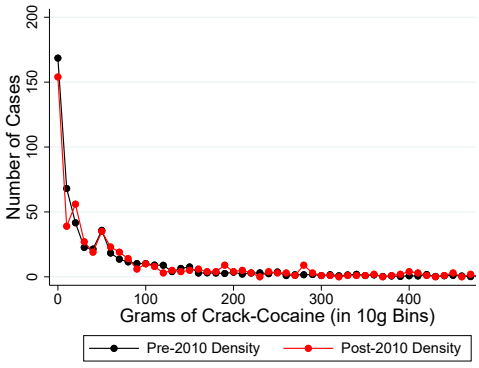
(a) All Offenders, Direct Comparison



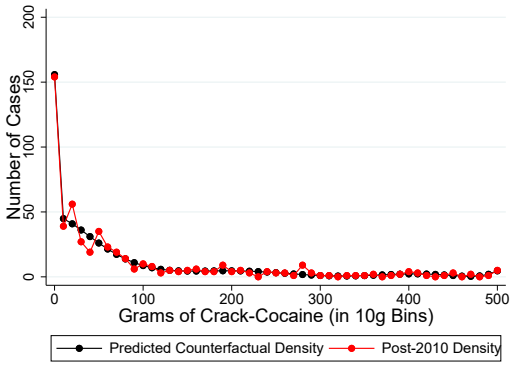
(b) All Offenders, Saez (2010) Method



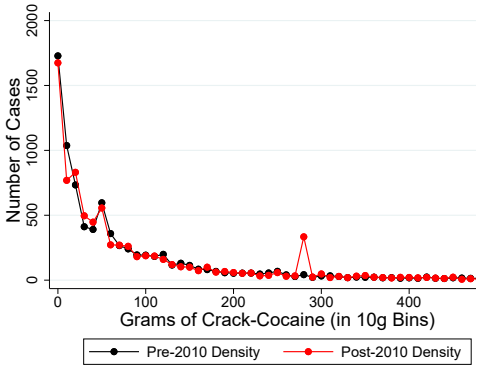
(c) White Offenders, Direct Comparison



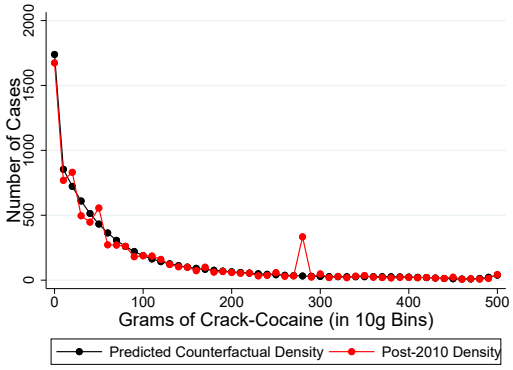
(d) White Offenders, Saez (2010) Method



(e) Black and Hispanic Offenders, Direct Comparison



(f) Black and Hispanic Offenders, Saez (2010) Method



Notes. Figure F1a plots the scaled density of drug quantities pre-2010 (in black) and the actual density of drug quantities post-2010 (in red) for all offenders. The amounts are aggregated into 10-gram bins and limited to drug quantities under 500g. Figures F1c and F1e do the same but restrict the sample to white offenders or black and Hispanic offenders, respectively. Figure F1b plots a predicted counterfactual density of drug quantities (in black) and the actual density of drug quantities post-2010 (in red). Figures F1d and F1f plot predicted counterfactual densities of drug quantities (in black) and the actual densities of drug quantities post-2010 (in red) by race. The amounts are aggregated into 10-gram bins and limited to drug quantities under 500g.

Table F1. All Bunching Results using Aggregated/Binned Comparison with Bootstrapped SEs

	Pr(280-290g Crack-Cocaine Recorded)					
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0333***		-0.0002***		-0.0006***	0.0772***
	(0.0017)		(0.0001)		(0.0002)	(0.0056)
After 2010 x White		0.0135**		-0.00008		
		(0.0057)		(0.0001)		
After 2010 x Non-White		0.0350***		-0.0002***		
		(0.0021)		(0.0001)		
Constant	-0.0003***	-0.0001***	0.000002	0.0000008***	0.000006***	-0.0008***
	(0.00002)	(0.0001)	(0.0000008)	(0.000001)	(0.000002)	(0.0001)
Data	USSC, Final Sentencing	USSC, Final Sentencing	NIBRS, Drug Seizures	NIBRS, Drug Seizures	DEA, Drug Seizures	EOUSA, Prosecutor Files
Bins	100	100	100	100	100	100
Observations	50,273	50,273	191,667	191,667	100,306	19,363

Notes. Bootstrapped standard errors in parentheses. Standard errors are calculated from the standard deviation in estimates derived from 50 replications where in each replication cases are sampled with replacement before aggregating to the 10g bin level. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. Specifications with the white/non-white and after 2010 interactions also include a dummy variable equal to one for black and Hispanic offenders. Columns 1-2 show the main bunching result for the final sentencing data. Columns 3-5 show no increase in bunching for drug seizure amounts. Column 6 shows an increase in bunching in prosecutor case management files.

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

Appendix G. Estimating the Conditional Disparity

In this section, I outline the assumptions necessary to estimate whether the racial disparity in bunching at 280g is due to differences in the underlying distributions of observed evidence or a difference in the likelihood a case is bunched **conditional** on the observed evidence. Conceptual discussion of these tests is in Section IV.B of the main text. This model closely follows Goncalves and Mello (2020).

I. Institutional Setting

Consider a simplified criminal court setting with drug cases, prosecutor discretion over amount charged, and mandatory minimum sentences. Assume the **seized evidence** s in a case is drawn from a discrete distribution $G_r(.)^t$ that is specific to each race r and time-period t (pre- vs. post-2010). The prosecutor for the case chooses the **amount (in grams) of drugs charged** a, and can charge amounts higher than s by collecting additional evidence $a-s$. Seized evidence s is a noisy measure of true drug trafficking.

I observe the amount charged a. Publicly available data from the USSC does not report the seized evidence s for each case, and true drug trafficking is unknown to the researcher and the prosecutor. The prosecutor chooses a based on a variety of factors. The first goal of this analysis is to identify racial disparities in a conditional on s (i.e. a **conditional racial disparity**). The second goal (addressed in Section V.D and in Appendix D) is to model under what conditions the disparity reflects discrimination by prosecutors and to conduct empirical tests of that model.

In this section, I detail the identifying assumptions necessary to estimate the conditional racial disparity. The set-up closely follows Goncalves and Mello (2020) who use a difference-in-bunching design to estimate police officer bias in speeding tickets. For now, consider the prosecutor's objective a function of tastes (including racial biases), career concerns, the sentence that would be justified under law if true drug trafficking were observed, and costs associated with building the case.

The amount of drugs charged a maps onto a **mandatory minimum sentencing schedule** $l(a)^t$ that differs pre-2010 $t=0$ and post-2010 $t=1$.

$$l(a)^t = \begin{cases} 1 & \text{if } a < mm_L^t \\ 5 & \text{if } mm_L^t \leq a < mm_U^t \\ 10 & \text{if } mm_U^t \leq a \end{cases} \tag{4}$$

If a is below the lower threshold for time period t, the defendant is sentenced to 1 year. If a is equal to or above the lower threshold but below the upper threshold, the defendant is sentenced to 5 years. If a is equal to or above the upper threshold, the defendant is sentenced to 10 years. A mandatory minimum does not, by law, require a discontinuous increase in sentence length at the thresholds. In practice, sentences do jump at 50g pre-2010 and 280g post-2010.

Given the seized evidence s (unobserved in the data but observed by the prosecutor) in the case and the defendant's race r (observed in the data), the prosecutor charges a final amount a (observed in the data) that is equal to a mandatory minimum threshold $mm = \{5, 28, 50, 280\}$ (i.e. "bunching" at the threshold) with a **bunching probability** $Pr(a = mm|s, r)^t$ (unobserved in the data). Finally, let defendants be in one of two broad **race** categories: white $r = w$ or black/Hispanic $r = bh$.

II. Defining the Conditional Racial Disparity

Now, I define a racial disparity in the amount charged a conditional on s and outline key equations.

There is a **conditional racial disparity** in bunching at 280g after 2010 if $Pr(a = 280|s, bh)^1 > Pr(a = 280|s, w)^1$. In other words, a conditional racial disparity exists if a black or Hispanic defendant with amount seized s is more likely to be bunched at 280g than a white defendant with the same amount seized s.

I observe the final amount charged, which can be written for the following ranges as:

$$Pr(a=j|r)^t = \begin{cases} \begin{cases} (a) & Pr(s=50|r)^0 + \sum_{k<50} Pr(s=k|r)^0 \times Pr(a=50|s=k,r)^0 & \text{if } j=50 \\ (b) & Pr(s=j|r)^0 & \text{if } 50<j \end{cases} & \text{if } t=0 \\ \begin{cases} (c) & Pr(s=j|r)^1 \times (1-Pr(a=280|s,r)^1) & \text{if } 50<j<280 \\ (d) & Pr(s=280|r)^1 + \sum_{k<280} Pr(s=k|r)^1 \times Pr(a=280|s=k,r)^1 & \text{if } j=280 \\ (e) & Pr(s=j|r)^1 & \text{if } 280<j \end{cases} & \text{if } t=1 \end{cases} \tag{5}$$

Equations **(5.a)** and **(5.b)** express the probability a case is charged with a given amount a prior to 2010. First, the probability a defendant is charged with an amount a equal to 50g is equal to the probability the seized evidence s is 50g plus the likelihood that a case with s under 50g gets moved up to 50g **(eqn. 5.a)**. Second, since there is no sentencing benefit of charging an amount above 50g, the probability a case is charged above 50g **(eqn 5.b)** is equal to the probability s is equal to that amount.

Equations **(5.c)-(5.e)** express the probability a case is charged with a given amount a after 2010. The probability a case is charged with an amount below 280g and above 50g **(eqn. 5.c)** is equal to the probability that s is equal to that amount and that the case does not get moved up to 280g given the amount s. The probability a case is charged with 280g **(eqn. 5.d)** is equal to the probability s is 280g plus the likelihood that a case with s under 280g gets moved up to 280g. As in **(eqn. 5.b)**, the probability a case is charged above 280g **(eqn 5.e)** is equal to the probability that s is equal to that amount. Throughout, I assume that prosecutors don't suppress evidence, i.e. $a \geq s$.[14]

III. Difference-in-Bunching Estimator and the Conditional Racial Disparity

To estimate whether $Pr(a=280|s,r)^1$ differs for black/Hispanic vs. white defendants, I compare the distribution of amounts charged after 2010 to the distribution of amounts charged prior to 2010.

Under the assumption that $Pr(s=k|r)^0 = Pr(s=k|r)^1$–i.e., the probability a case with a defendant of race r has seized evidence $s=k$ does not change pre- vs. post-2010–the **difference-in-bunching coefficients (eqn. 2 of main text)** $\delta-\beta$ yields the following:

$$\delta-\beta = [\sum_{k<280} Pr(s=k|bh) \times Pr(a=280|s=k,bh)^1] - [\sum_{k<280} Pr(s=k|w) \times Pr(a=280|s=k,w)^1] \tag{6}$$

$\delta>0$ and $\beta>0$ imply that prosecutors increase a in response to the Fair Sentencing Act, and $\delta-\beta>0$ implies that they increase a more for black and Hispanic defendants. This alone is of interest–it shows that prosecutors use their discretion to increase sentences in response to the FSA and that the burden of this falls on minority defendants. However, $\delta-\beta>0$ could be driven by different underlying distributions of seized evidence s (i.e. different $Pr(s=k|r)$) or by disparate treatment conditional on s (i.e. different $Pr(a=280|s,r)^1$–a conditional racial disparity).

The goal of this section is to outline how to test whether $\delta-\beta>0$ is due to a conditional racial disparity. I detail two tests. For the first test, $\delta-\beta$ can be rewritten as follows:

$$\delta-\beta = \overbrace{[\sum_{k\leq 50} Pr(s=k|bh) \times Pr(a=280|s=k,bh)^1 - \sum_{k\leq 50} Pr(s=k|w) \times Pr(a=280|s=k,w)^1]}^{H} + \overbrace{[\sum_{50<k<280} Pr(s=k|bh) \times Pr(a=280|s=k,bh)^1 - \sum_{50<k<280} Pr(s=k|w) \times Pr(a=280|s=k,w)^1]}^{I} \tag{7}$$

[14] In reality, it is possible for prosecutors to reduce the drug amount charged or choose not to pursue a drug charge entirely. Introducing this possibility means the disparity in bunching could be due to: (1) a difference in underlying observed drugs, (2) a conditional disparity in bunching, or (3) a conditional disparity in suppressing. The empirical evidence I show is consistent with (2) and (3), both of which are disparities conditional on underlying observed drugs. For that reason, I focus on the simpler case.

First, I test whether the H term can explain $\delta - \beta > 0$. I observe $Pr(a = 50|r)^0$ and $Pr(a = 50|r)^1$. Equation **(5)** implies that:

$$Pr(a = 50|bh)^1 - Pr(a = 50|bh)^0 = -[Pr(s = 50|bh) \times Pr(a = 280|s = 50, bh)^1] - [\sum_{k<50} Pr(s = 50|bh) \times Pr(a = 50|s = k, bh)^0] \quad (8)$$

Under the assumption that $Pr(a = 50|s, r)^0 \geq Pr(a = 280|s, r)^1$ for all $s \leq 50$, equation **(8)** is greater than the $\sum_{k \leq 50} Pr(s = k|bh) \times Pr(a = 280|s = k, bh)^1$ term from equation **(7)**. Thus, if the sum of equation **(8)** and $\delta - \beta$ is greater than zero, then the term H cannot explain $\delta - \beta > 0$. In other words, the shift **from** 50g for black and Hispanic offenders is an upper bound for the movement **to** 280g that can be explained by amounts seized at 50g or below. If this shift is not enough to explain the racial disparity in bunching at 280g, then the racial disparity must be due to term I. This test is in paragraph 2 of Section V.A.3.

Second, I test whether racial differences in $\sum_{50<k<280} Pr(s = k|r)$ from term I can explain $\delta - \beta > 0$. From equation **(5.b)**, $Pr(a = k|r)^0 = Pr(s = k|r)^0 \; \forall \; 280 > k > 50$. Thus, I can test if $\sum_{50<k<280} Pr(s = k|w)^0 = \sum_{50<k<280} Pr(s = k|bh)^0$ by testing if $\sum_{50<k<280} Pr(a = k|w)^0 = \sum_{50<k<280} Pr(a = k|bh)^0$. In other words, if the distributions of pre-2010 charged amounts from 50-280g are approximately equal by race, then the racial disparity in bunching must be due to a racial disparity in the probability a case is bunched at 280g **conditional** on the seized evidence. This test is in Figure 2a.

Now, I turn to the second test for a conditional racial disparity. The assumptions above also imply:

$$Pr(a = 50 < k < 280|r)^1 = Pr(s = k|r)^1 \times (1 - Pr(a = 280|s = k, r)^1) \quad (9)$$

$$Pr(a = 50 < k < 280|r)^0 = Pr(s = k|r)^0 \quad (10)$$

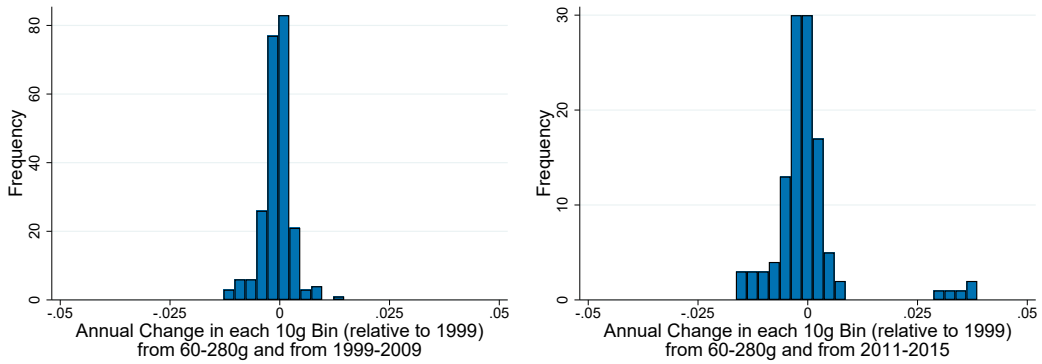
The difference between equation **(9)** and **(10)** by race can be estimated as follows:

$$(\text{Charged X-Yg})_{it} = \alpha + \delta^X(\text{After2010}\times\text{BlackOrHispanic})_{it} + \gamma\text{After2010}_{it} + \lambda\text{BlackOrHispanic}_i + \epsilon_{it} \quad (11)$$

The coefficient $\delta^X = Pr(a = 280|w, s)^1 - Pr(a = 280|bh, s)^1$. Then, $\delta^X < 0$–i.e., black and Hispanic defendants are more likely to be shifted away from a given amount X after 2010–implies that there is a racial disparity in amount charged a conditional on the underlying evidence seized s. This test is in Figure 2b.

Figures G1 and G2 provide support for the assumptions that the pre-2010 distribution from 60-280g is a good counterfactual. One concern is that drug involvement is changing over time such that involvement in one set of years will differ from involvement in other years even in the absence of a policy change. Figure G1 shows that the changes (relative to 1999) in amounts from 60-280g (including 280-290g) are much larger and more disperse after the FSA than in the years before the FSA. Before the FSA, changes are centered tightly around zero. Figure G2 focuses on the five ranges that show marked racial disparities in their decreases after 2010. The event study plots suggest that these ranges did not exhibit much change from 1999-2010, and that the decreases began specifically after the FSA.

Figure G1. Histograms of Changes in Each 10g Bin from 60-280g (Including 280-290g)

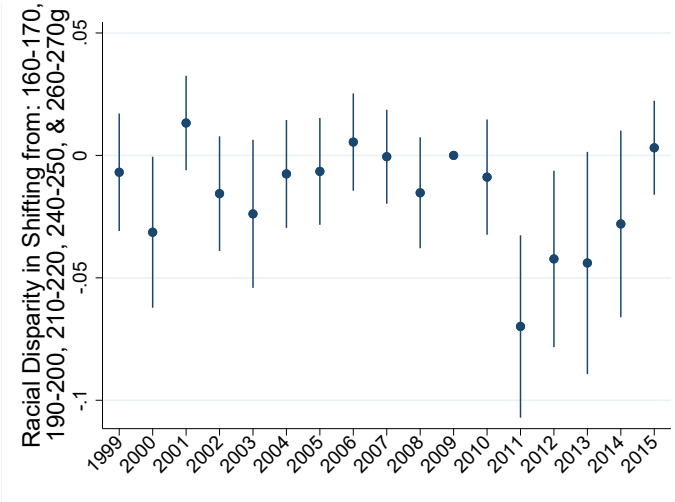


Notes. Both panels plot a histogram of coefficient estimates from the following regression:

$$(\text{Charged X-Yg})_{it} = \sum_{t=2000}^{Y} \text{Year}_t + \text{Trend}_t + \epsilon_{it}$$

where the summation represents a set of year dummy variables from the year 2000 to the year Y. The trend variable is a linear trend in year. The dependent variable is equal to one if the case is recorded with an amount in the 10g bin from X-Yg, where I consider all bins from 60-280g, including 280-290g. In panel (a), Y=2009. Panel (a) thus plots the coefficients on the year dummy variables from 2000-2009 with the year 1999 as the reference category. The estimated changes in the probabilities of being recorded in each bin do not vary much from year to year before 2010. These are mostly centered around zero. In panel (b), Y=2015. Panel (b) plots the coefficients on the year dummy variables from 2011-2015 with the year 1999 as the reference category. Still, many of the changes are near zero. However, there is a dramatic shift to the right (for the bin 280-290g) and a shift to the left (for bins above 60g but below 280g). In general, 60% of the estimated changes (for bins from 60-280g, excluding 280-290g) pre-2010 are below zero, compared to almost 70% post-2010. 29.5% of changes pre-2010 are below -0.002 compared to 43.6% of changes post-2010. The median change pre-2010 is -0.0006 and the median change post-2010 is -0.0016. The 10th percentile change pre-2010 is -0.0039 and the 10th percentile change post-2010 is -0.0085. These figures are based on the USSC data. See Table 1 for notes about data construction.

Figure G2. Event Study for Changes in Bins with Large Decreases After 2010



Notes. Figure 2b indicates that there are racial disparities in the movement away from the following bins after 2010: 160-170g, 190-200g, 210-220g, 240-250g, and 260-270g. This figure plots event study coefficients of the disparity in those bins for each year relative to 2009. The differences by race are relatively small from 1999-2010 with decreases in the probability of being in those bins for black and Hispanic offenders after 2010. This figure is based on the USSC data. See Table 1 for notes about data construction.

Appendix References

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

Best, M., J. Cloyne, E. Ilzetzki, and H. J. Kleven. 2020. "Estimating the Elasticity of Intertemporal Substitution Using Mortgage Notches." *Review of Economic Studies* 87 (2): 656-690.

Bjerk, D. 2017. "Mandatory Minimums and the Sentencing of Federal Drug Crimes." *Journal of Legal Studies* 46 (1): 93-128

BJS. "Survey of Inmates in State and Federal Correctional Facilities, 2004." https://www.icpsr.umich.edu/web/NACJD/studies/4572.

Brown, K. M. 2013. "The Link between Pensions and Retirement Timing: Lessons from California Teachers." *Journal of Public Economics* 98 (February): 1–14.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik. 2017. "rdrobust: Software for Regression Discontinuity Designs." *Stata Journal*, 17(2): 372-404.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer. 2019. "The Effect of Minimum Wages on Low-Wage Jobs: Evidence from the United States Using a Bunching Estimator." *Quarterly Journal of Economics* 134 (3): 1405-1454.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics* 126 (2): 749–804.

Craigie, T. and M. Zapryanova. 2020. "Mandatory Minimum Reforms, Sentencing, and Racial-Ethnic Disparities." Working paper.

Dale v. United States, 139 S.Ct. 287, Petition for Writ of Certiorari. 2018. https://www.supremecourt.gov/DocketPDF/18/18-5663/60026/20180820134819671_00000007.pdf.

DEA. "System to Retrieve Information from Drug Evidence: Cocaine Records, 1999-2015." Obtained from FOIA request made here: https://www.dea.gov/divisions/requesting-dea-foia-records.

Didwania, S. 2020. "Mandatory Minimums and Federal Sentences." Working paper.

EOUSA. "National Caseload Data, September 2017 data files." https://www.justice.gov/usao/resources/foia-library/national-caseload-data (link resolves to current EOUSA files).

EOUSA. 2016. "Legal Information Office Network System User's Manual: Release 5.4.2." https://www.justice.gov/usao/file/835096/download.

Fack, G., and C. Landais. 2016. "The Effect of Tax Enforcement on Tax Elasticities: Evidence from Charitable Contributions in France." *Journal of Public Economics* 133 (January): 23–40.

Feigenberg, B. and C. Miller. 2020. "Racial Divisions and Criminal Justice: Evidence from Southern State Courts." Forthcoming at *American Economic Journal: Economic Policy*.

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.

FL Department of Corrections. "OBIS Database, April 2016 version." http://www.dc.state.fl.us/pub/obis_request.html (link resolves to current OBIS database).

Flowers v. Mississippi, 588 U.S. __. 2019. https://www.law.cornell.edu/supremecourt/text/17-9572.

Free Law Project. "Court Listener: Advanced RECAP Search." https://www.courtlistener.com/recap/.

Gelber, A. M., D. Jones, and D. W. Sacks. 2020. "Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test." *American Economic Journal: Applied Economics* 12 (1): 1-31.

Goncalves, F. and S. Mello. 2020. "A Few Bad Apples? Racial Bias in Policing." Working paper.

Honold, D. 2014. "Note: Quantity, Role, and Culpability in the Federal Sentencing Guidelines." *Harvard Journal on Legislation* 51: 389-314.

Jones v. United States, 526 U.S. 227. 1999. https://www.law.cornell.edu/supct/html/97-6203.ZO.html.

judyrecords. "464 million United States Court Cases." https://www.judyrecords.com/.

Kleven, H. J. 2016. "Bunching." *Annual Review of Economics* 8: 435–464.

Lu, L. D. 2007. "Prosecutorial Discretion and Racial Disparities in Federal Sentencing: Some Views of Former U.S. Attorneys." *Federal Sentencing Reporter* 19 (3): 192-201.

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

National Registry of Exonerations. "Browse Cases: Detailed View." http://www.law.umich.edu/special/exoneration/Pages/detaillist.aspx.

NORC at the University of Chicago. "General Social Survey, 1972-2018." https://gss.norc.org/.

O'Neil v. United States, 137 S.Ct. 223. 2013. https://www.leagle.com/decision/insco20131007q42.

Possley, M. and K. Armstrong. 1999. "Part 2: The flip side of a fair trial." *Chicago Tribune* (Chicago, IL).

Prasad, P. 2018. "Implicit Racial Biases in Prosecutorial Summations: Proposing an Integrated Response." *Fordham Law Review* 86 (6): 3091-3126.

Project Implicit. "Race IAT, 2006-2016." https://osf.io/52qxl/.

Rehavi, M. M., and S. B. Starr. 2013. "Mandatory Sentencing and Racial Disparity: Assessing the Role of Prosecutors and the Effects of Booker." *The Yale Law Journal* 123 (2): 2-80.

Rehavi, M. M., and S. B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6):1320–54.

Rhodes, W., R. Kling, J. Luallen, and C. Dyous. 2015. "Federal Sentencing Disparity: 2005-2012." Bureau of Justice Statistics Working Paper Series.

Rozema, K. 2020. "Occupational Licensing and Legal Services: Evidence from Diploma Privileges and Lawyer Sanctions." Working paper.

Saez, E. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3): 180–212.

Stephens-Davidowitz, S. "Racial Animus (Google Search Data): IAT, Google, and Obama." http://sethsd.com/s/StateIATGoogle-4say.csv.

United States v. Andre Johnson, No. 03-6477. 6th Cir. 2005. https://www.courtlistener.com/opinion/2972399/united-states-v-johnson/.

United States v. Dana Cooper, No. 4:15-cr-20155. E.D. Michigan. 2015. https://www.courtlistener.com/docket/16819229/11/united-states-v-cooper/.

United States v. David Tyrone Murrell, No. 7:12-cr-00014-D All Defendants. E.D. North Carolina. 2017. https://www.courtlistener.com/docket/5634402/126/united-states-v-murrell/.

United States v. Eric Lanard Williams, No. 14-11943. 11th Cir. 2015. https://www.judyrecords.com/record/5f3a2cmyzb996.

United States v. Hector Morales, No. 15-438-cr. 2d Cir. 2018. https://www.courtlistener.com/opinion/4492185/united-states-v-morales/.
United States v. Kizzy Fader, No. 1:12-cr-00007-JAW. D. Maine. 2016. https://www.courtlistener.com/docket/5109479/117/united-states-v-fader/.
United States v. Karen Antoinette Johnson, No. 16-3280. 10th Cir. 2017. https://www.judyrecords.com/record/5d95m1mzhe830.
United States v. Marvin Lee Ellis, Nos. 14-3165 & 14-3181. 10th Cir. 2017. https://www.judyrecords.com/record/50q3p0b6se00b.
USSC. 2015. "United States Sentencing Commission (USSC) Guidelines Manual." https://www.ussc.gov/guidelines/guidelines-archive/2015-guidelines-manual.
Substance Abuse and Mental Health Data Archive. "National Survey on Drug Use and Health, 2002-2016." https://www.datafiles.samhsa.gov/study/national-survey-drug-use-and-health-nsduh-2002-2016-nid18454.
US DOJ, FBI. "National Incident-Based Reporting System (NIBRS) Series, 2000-2015." https://www.icpsr.umich.edu/icpsrweb/NACJD/series/128.
USSC. "Monitoring of Federal Criminal Sentences Series, 1999-2016." https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/83.
USSC. "Monitoring of Federal Criminal Convictions and Sentences: Appeals Data Series, 2012-2013." https://www.icpsr.umich.edu/web/NACJD/series/75.
Webb, D. K. and S. F. Turow. 1982. "The Prosecutor's Function in Sentencing." *Loyola U. Chicago Law Journal* 13 (4): 641-668.
Wright, R. 2006. "Federal or State? Sorting as a Sentencing Choice." *Criminal Justice* 21 (2): 16-21.
Yang, C. "Replication data for: Resource Constraints and the Criminal Justice System: Evidence from Judicial Vacancies." https://test.openicpsr.org/openicpsr/project/114590/version/V1/view?path=/openicpsr/114590/fcr:versions/V1/Data_2015_0150/Public-Use-Data&type=folder.
Zaresani, A. 2020. "Adjustment Costs and Incentives to Work: Evidence from a Disability Insurance Program." *Journal of Public Economics* 188.