Appendix A. Additional Analyses

Table A1. Summary Statistics for FL, NIBRS, and DEA Records

	Pre-2010	Post-2010	Observations
Panel A. Cocaine Felony Convictions in FL			
200-400g	0.00474	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)	10.33	7.76	203,700
	(46.19)	(44.87)	
280-290g	0.000360	0.000141	203,700
	(0.0190)	(0.0119)	
Black	0.737	0.746	191,774
	(0.440)	(0.435)	
Male	0.837	0.834	192,721
	(0.370)	(0.372)	
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	0.544	100,302
	(0.499)	(0.498)	
Price per gram (median)	42.02	47.62	37,820

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 A2. 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.

Table A3. Result Robust to Other Drug Weight Sample Restrictions

	Pr(280-290g Crack-Cocaine)				
	(1)	(2)	(3)	(4)	(5)
After 2010 x White	0.0119**	0.0115**	0.0115**	0.0844***	0.0258**
	(0.0050)	(0.0049)	(0.0049)	(0.0131)	(0.0116)
After 2010 x Black or Hispanic	0.0345***	0.0329***	0.0328***	0.1186***	0.0718***
	(0.0021)	(0.0020)	(0.0020)	(0.0040)	(0.0042)
Constant	0.0031***	0.0030***	0.0030***	0.0034***	0.0088***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0027)
P-value: W = BH	0.0000	0.0000	0.0001	0.0127	0.0002
Sample Restriction	0-2500g	0-25000g	No Restriction	0-1000g	50-1000g
Includes Weights Coded as a Range	No	No	No	Yes	No
Observations	55,729	58,116	58,645	59,677	24,905

Notes. Robust standard errors in parentheses. 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 outliers 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 excludes drug weights below 50g (i.e. excluding weights close to the 5-year mandatory minimum pre- and post-2010).
*** 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	0.0314***		0.0336***		0.0304***	
	(0.0021)		(0.0021)		(0.0022)	
After 2010 x White		0.0125**		0.0128**		0.0128**
		(0.0053)		(0.0054)		(0.0054)
After 2010 x Black or Hispanic		0.0327***		0.0348***		0.0317***
		(0.0022)		(0.0022)		(0.0023)
Constant	0.0053***	0.0032***	0.0062***	0.0030**	0.0063***	0.0030**
	(0.0004)	(0.0010)	(0.0006)	(0.0015)	(0.0006)	(0.0015)
P-value: W = BH	-	0.0004	-	0.0002	-	0.0013
Hispanic Offenders Excluded	Yes	Yes	No	No	Yes	Yes
Post-2006 Data Only	No	No	Yes	Yes	Yes	Yes
Observations	47,763	47,763	25,893	25,846	23,241	23,241

Notes. Robust standard errors in parentheses. 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 Other Categorizations of Bunching

	Pr(280-300g)	Pr(280-320g)	Pr(280-380g)
	(1)	(2)	(3)
After 2010 x White	0.0154**	0.0146**	0.0137*
	(0.0061)	(0.0067)	(0.0083)
After 2010 x Black or Hispanic	0.0360***	0.0367***	0.0394***
	(0.0022)	(0.0025)	(0.0029)
Constant	0.0055***	0.0099***	0.0230***
	(0.0013)	(0.0017)	(0.0026)
P-value: W = BH	0.0016	0.0019	0.0033
Observations	52,745	52,745	52,745

Notes. Robust standard errors in parentheses. 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." Each column corresponds to a different definition 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 1-3, I use alternative ranges: 280-300g, 280-320g, and 280-380g.
*** p<0.01, ** p<0.05, * p<0.1

Table A6. 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	0.0347***		0.0348***		0.0345***		0.0327***		0.0322***	
	(0.0082)		(0.0081)		(0.0079)		(0.0068)		(0.0066)	
After 2010 x White		0.0125**		0.0130**		0.0136**		0.0118*		0.0138**
		(0.0058)		(0.0059)		(0.0062)		(0.0060)		(0.0066)
After 2010 x Black or Hispanic		0.0360***		0.0363***		0.0358***		0.0340***		0.0333***
		(0.0086)		(0.0086)		(0.0084)		(0.0073)		(0.0071)
Constant	0.0051***	0.0032***	0.0085***	0.0064**	0.0088**	0.0085**	0.0078*	0.0074*	0.0082**	0.0075**
	(0.0005)	(0.0010)	(0.0029)	(0.0031)	(0.0035)	(0.0037)	(0.0043)	(0.0044)	(0.0031)	(0.0033)
P-value: W = BH	-	0.0181	-	0.0184	-	0.0282	-	0.0286	-	0.0695
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	56,826	52,692	51,813	51,746	51,813	51,746	51,804	51,737	51,804	51,737

Notes. Standard errors clustered at the state-level in parentheses. 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, 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 year trend 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 by race.
*** p<0.01, ** p<0.05, * p<0.1

Table A7. 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	0.5747***	0.5606***	0.2046*	1.6031***	1.4119***	0.4804*	1.1208***	1.1615***	0.1102	0.0125**	0.0252**	0.0137*
	(0.1651)	(0.1840)	(0.1085)	(0.4518)	(0.4546)	(0.2498)	(0.4042)	(0.2758)	(0.5745)	(0.0053)	(0.0113)	(0.0083)
After 2010 x BH	0.8159***	0.9008***	0.3851***	2.0784***	2.0895***	0.8400***	2.1129***	2.1042***	0.8604	0.0360***	0.0710***	0.0394***
	(0.0337)	(0.0374)	(0.0235)	(0.0869)	(0.0878)	(0.0500)	(0.3645)	(0.2726)	(0.6351)	(0.0021)	(0.0042)	(0.0029)
Constant	-2.7258***	-2.3912***	-1.9948***	-5.7392***	-4.7715***	-3.7476***	3.5423***	2.6237***	3.6109***	0.0032***	0.0084***	0.0230***
	(0.0994)	(0.1102)	(0.0470)	(0.3020)	(0.3028)	(0.1138)	(0.3624)	(0.2202)	(0.3624)	(0.0010)	(0.0025)	(0.0026)
P-value: W = BH	0.1524	0.0701	0.1041	0.3015	0.1433	0.1580	0.0157	0.0007	0.3286	0.0000	0.0001	0.0033
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	52,745	25,647	52,745	52,745	25,647	52,745	400	380	400	52,745	25,647	52,745

Notes. Robust standard errors in parentheses. 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 5 of Table A3). Columns 3, 6, 9, and 12 extend the "bunching" definition to 280-380g (following column 3 of Table A5).

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

Table A8. 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***	0.0242***	0.0005	0.0727***
	(0.0087)	(0.0059)	(0.0003)	(0.0032)
After 2010 x Black or Hispanic	0.0833***	0.0441***	0.0093***	0.2030***
	(0.0027)	(0.0021)	(0.0006)	(0.0031)
Constant	0.0033***	0.0024***	0.0004***	0.8680***
	(0.0009)	(0.0007)	(0.0001)	(0.0021)
P-value: W = BH	0.0063	0.0016	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	67,040	65,003	149,428	149,428

Notes. Robust standard errors in parentheses. 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 missing exact weights (i.e. 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. 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 A9. Difference-in-Difference Bunching Identification

	Pr(280-290g)				Pr(50-60g)	
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0011*	-0.0002				
	(0.0006)	(0.0011)				
After 2010 x Crack-cocaine	0.0336***	0.0127**				
	(0.0021)	(0.0054)				
After 2010 x Crack-cocaine x Black or Hispanic		0.0217***				
		(0.0059)				
Crack-cocaine	-0.0020***	-0.0042***	-0.0036**	0.0088	0.0151***	0.0210*
	(0.0005)	(0.0011)	(0.0016)	(0.0058)	(0.0053)	(0.0122)
Crack-cocaine x Black or Hispanic			0.0020	0.0229***	0.0108*	-0.0021
			(0.0017)	(0.0063)	(0.0057)	(0.0127)
Constant	0.0072***	0.0074***	0.0068***	0.0070***	0.0502***	0.0438***
	(0.0003)	(0.0006)	(0.0012)	(0.0026)	(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	149,428	149,428	65,475	17,307	65,475	17,307

Notes. Robust standard errors in parentheses. 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.

*** 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	-0.0030	-0.1162***	0.0326**	-0.0006	0.0189
	(0.0179)	(0.0188)	(0.0149)	(0.0111)	(0.0143)
After 2010 x Black or Hispanic	0.0222***	-0.0696***	0.0341***	-0.0066**	-0.0100***
	(0.0039)	(0.0050)	(0.0041)	(0.0029)	(0.0038)
Constant	0.1971***	0.3242***	0.0968***	0.0653***	0.0965***
	(0.0068)	(0.0080)	(0.0050)	(0.0042)	(0.0050)
P-value: W = BH	0.1669	0.0164	0.9216	0.6000	0.0503
Observations	52,745	52,745	52,745	52,745	52,745
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)
	(1)	(2)	(3)	(4)	(5)
After 2010 x White	0.0028	0.0125**	0.0137	0.0099	0.0294***
	(0.0162)	(0.0053)	(0.0096)	(0.0070)	(0.0090)
After 2010 x Black or Hispanic	-0.0165***	0.0360***	0.0044*	0.0016	0.0044**
	(0.0045)	(0.0021)	(0.0025)	(0.0018)	(0.0020)
Constant	0.1493***	0.0032***	0.0353***	0.0163***	0.0160***
	(0.0061)	(0.0010)	(0.0032)	(0.0022)	(0.0021)
P-value: W = BH	0.2503	0.0000	0.3470	0.2539	0.0066
Observations	52,745	52,745	52,745	52,745	52,745

Notes. Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. 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."
*** p<0.01, ** p<0.05, * p<0.1

Table A10b. Missing Mass in the Distribution of Drug Amounts by Race, with Various Time Trend Controls and State FEs

	Pr(< 280g)	Pr(280-290g)	Pr(> 290g)
	(1)	(2)	(3)
Panel A. No Interaction with Time Trend			
After 2010 x White	-0.0685***	0.0120**	0.0566***
	(0.0151)	(0.0055)	(0.0143)
After 2010 x Black or Hispanic	-0.0602***	0.0343***	0.0259***
	(0.0051)	(0.0023)	(0.0047)
Constant	0.9372***	0.0059***	0.0569***
	(0.0053)	(0.0013)	(0.0052)
P-value: W = BH	0.5840	0.0001	0.0330
Observations	52,678	52,678	52,678
Panel B. Interaction with Linear Time Trend			
After 2010 x White	-0.0403*	0.0164**	0.0240
	(0.0229)	(0.0083)	(0.0218)
After 2010 x Black or Hispanic	-0.0601***	0.0345***	0.0256***
	(0.0064)	(0.0033)	(0.0057)
Constant	0.9078***	0.0043**	0.0880***
	(0.0100)	(0.0020)	(0.0098)
P-value: W = BH	0.4063	0.0418	0.9418
Observations	52,678	52,678	52,678
Panel C. Interaction with Quadratic Time Trends			
After 2010 x White	0.0031	0.0133	-0.0164
	(0.0303)	(0.0099)	(0.0291)
After 2010 x Black or Hispanic	-0.0256***	0.0301***	-0.0045
	(0.0085)	(0.0040)	(0.0078)
Constant	0.8789***	0.0038	0.1173***
	(0.0192)	(0.0040)	(0.0188)
P-value: W = BH	0.3614	0.1150	0.6933
Observations	52,678	52,678	52,678

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 1 for notes about sample selection. 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 general model I estimate is:

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

$Trend$ takes on the value of zero (i.e. no trend interaction), a linear trend, or a quadratic trend. $g(t)_t$ is a linear trend when no trend interactions are used and when the linear trend interaction is used. $g(t)_t$ is a quadratic trend when the quadratic trend interactions are used. Figures A7j-k show the total share of cases below 280g and above 280g over time, by race. For these shares, there are considerable trends over time, especially for white offenders. To quantify the break in those trends after 2010, I estimate case-level regressions that interact the dummy variable for after 2010 with a linear time trend centered at zero in 2011. Panel (a) shows the estimates without accounting for these time trends, and as a result, column 3 indicates that white offenders are more likely to be charged with amounts greater than 290g after 2010, relative to black and Hispanic offenders. This is true, but it is due to a substantial rise in cases above 290g for white offenders that begins in 2005. Panels (b) and (c) account for this by estimating the break in the trend after 2010. Both panels indicate that white, black, and Hispanic offenders have similar (and small) trend breaks in their share of cases above 290g. Likewise, both panels show bunching at 280-290g, a racial disparity in bunching, and evidence that the excess mass at 280-290g is drawn from cases that would have been charged below 280g prior to 2010. All specifications include state fixed-effects (Z_i).
*** p<0.01, ** p<0.05, * p<0.1

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

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 (Actual Change)	0.0246***	-0.0710***	0.0323***	-0.0098***	-0.0120***
	(0.0042)	(0.0055)	(0.0044)	(0.0033)	(0.0042)
Constant	0.1065***	0.2920***	0.1134***	0.0751***	0.1263***
	(0.0024)	(0.0035)	(0.0025)	(0.0021)	(0.0026)
Predicted Change from Conceptual Model	Increase	Decrease	Increase	Decrease	Decrease
Observations	25,893	25,893	25,893	25,893	25,893
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)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	-0.0108**	0.0336***	0.0050*	0.0026	0.0056**
	(0.0050)	(0.0021)	(0.0027)	(0.0019)	(0.0022)
Constant	0.1886***	0.0062***	0.0443***	0.0207***	0.0269***
	(0.0031)	(0.0006)	(0.0016)	(0.0011)	(0.0013)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	25,893	25,893	25,893	25,893	25,893

Notes. Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams and sentenced from 2007-2015. 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."
*** p<0.01, ** p<0.05, * p<0.1

Table A10d. Missing Mass in the Distribution of Drug Amounts, Trial Cases Only

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 (Actual Change)	0.0294	-0.0530**	0.0248	-0.0120	-0.0591***
	(0.0181)	(0.0216)	(0.0171)	(0.0137)	(0.0126)
Constant	0.1104***	0.2592***	0.0984***	0.0831***	0.1112***
	(0.0064)	(0.0089)	(0.0061)	(0.0056)	(0.0064)
Predicted Change from Conceptual Model	Increase	Decrease	Increase	Decrease	Decrease
Observations	2,841	2,841	2,841	2,841	2,841
R-squared	0.030	0.020	0.006	0.008	0.007
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)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	-0.0392**	0.0749***	0.0030	0.0217*	0.0085
	(0.0199)	(0.0131)	(0.0124)	(0.0111)	(0.0111)
Constant	0.2050***	0.0033***	0.0562***	0.0281***	0.0389***
	(0.0082)	(0.0012)	(0.0047)	(0.0034)	(0.0039)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	2,841	2,841	2,841	2,841	2,841
R-squared	0.012	0.022	0.007	0.006	0.010

Notes. Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams and 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."
*** p<0.01, ** p<0.05, * p<0.1

Table A11. Sentencing Consequences of Being Above the Threshold Amount

	Years Sentenced					
	(1)	(2)	(3)	(4)	(5)	(6)
Above 280g	-0.580**	0.0621			0.00410	-0.0576
	(0.289)	(0.691)			(0.294)	(0.461)
Above 280g x After 2010	2.332***	2.181**			0.971*	2.836***
	(0.508)	(1.102)			(0.535)	(0.842)
Above 50g			0.755***	0.955***	1.469***	2.101***
			(0.128)	(0.158)	(0.180)	(0.227)
Above 50g x After 2010			-1.387***	-1.063***	-1.298***	-2.058***
			(0.270)	(0.357)	(0.451)	(0.445)
Constant	12.93***	11.48***	9.664***	9.540***	13.12***	14.08***
	(0.170)	(0.565)	(0.114)	(0.116)	(3.298)	(3.709)
Bandwidth	±250g	±50g	±250g	±50g	±250g	±250g
Includes Life & <1 Month	No	No	No	No	No	Yes
Observations	29,767	2,800	49,154	14,713	29,064	31,134
R-squared	0.037	0.015	0.070	0.035	0.038	0.031

Notes. Robust standard errors in parentheses. The estimates in this table are based on the USSC data. 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 (4) 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 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 3a-d show graphical evidence of the sentencing penalty. Figure A9 shows that the estimate of the sentencing penalty from model (5) is robust to many different bandwidths from 10g to 250g.
*** 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(280-290g)	Pr(280-290g)
	(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	-	0.2382
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	-0.000186**		0.0332**	
	(8.67e-05)		(0.0162)	
After 2010 x White		0.0002		0.0038
		(0.0002)		(0.0513)
After 2010 x Black or Hispanic		-0.0003***		0.0346**
		(0.0001)		(0.0164)
Constant	0.000422***	0.0003***	0.143***	0.1558***
	(4.94e-05)	(0.0001)	(0.0120)	(0.0219)
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	-	0.0830	-	0.5469
Observations	219,515	219,515	4,376	4,376

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 (1)	280-290g (2)	280-290g (3)	Weight (g) (4)
After 2010	0.0826***	0.0760***	0.0989***	26.09***
	(0.0180)	(0.0191)	(0.0129)	(5.659)
After 2010 × Any Federal	-0.00889			
	(0.0190)			
After 2010 × Only Federal		-0.00263		
		(0.0202)		
After 2010 × FBI			0.0160	52.99***
			(0.0198)	(11.29)
After 2010 × ATF			-0.0732***	-15.03**
			(0.0143)	(6.953)
After 2010 × State/local			-0.0229	-7.648
			(0.0231)	(11.45)
After 2010 × DEA & State/local			-0.0133	-3.980
			(0.0383)	(19.46)
After 2010 × Joint state/local			0.0148	7.345
			(0.0507)	(25.60)
After 2010 × ATF & State/local			-0.00860	-9.386
			(0.0388)	(13.18)
After 2010 × FBI & State/local			-0.0619	-17.32
			(0.0386)	(22.44)
Constant	0.00342***	0.00360***	0.00481***	77.73***
	(0.00121)	(0.00136)	(0.000876)	(1.523)
Observations	17,042	15,016	17,042	17,042

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 1-3 show the relationship between race of offender and drug weight seized, in general. Column 4 shows how the weight of an offender's seized drugs changes by race after 2010. Columns 5-11 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. 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 sentence year-month. I code "% Black or Hispanic" equal to the fraction of offenders 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-month. The number of observations falls because not all cases that enter EOUSA end in a sentence. Coefficients are estimated from the following regression for columns 2 and 3 (with only the dependent variable changing):

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

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

Table A16. 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)	
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 A17. 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	-0.2193***	0.2421***	-0.0228	-0.1143	0.1882***	-0.0739
(15+ cases post-2010)	(0.0459)	(0.0339)	(0.0272)	(0.0806)	(0.0447)	(0.0640)
Constant	0.9309***	-	0.0691***	0.8855***	-	0.1145*
	(0.0242)	-	(0.0242)	(0.0617)	-	(0.0617)
Observations	1,647	1,647	1,647	699	699	699
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, column (2) is a mechanical relationship, hence the missing standard error. 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 A18. Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors, Leave-One-Out Classification

Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)
Atty. Bunches at 280-290g Post-2010	-0.114*	0.149***	-0.0354
(Leaving out current case in calculation)	(0.0659)	(0.0435)	(0.0463)
Constant	0.891***	0.0272***	0.0816*
	(0.0432)	(0.00765)	(0.0436)
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 280-290g	-0.505***	0.527***	-0.0227
(Leaving out current case in calculation)	(0.116)	(0.0717)	(0.0976)
Constant	0.891***	0.0380***	0.0708**
	(0.0346)	(0.00791)	(0.0349)
Observations	971	971	971

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). **The classification for each bunching attorney is based on all cases excluding the current observation (i.e. a leave-one-out procedure).** Coefficients in panel B 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.
*** p<0.01, ** p<0.05, * p<0.1

Table A19. Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors, with Bootstrapped SEs

Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)
Atty. Bunches at 280-290g Post-2010	-0.1794***	0.2170***	-0.0376
	(0.0659)	(0.0371)	(0.0510)
Constant	0.9184***	-	0.0816*
	(0.0435)	-	(0.0435)
Observations	989	989	989
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Atty. Bunches at 50-60g Pre-2010	-0.0785***	0.0575***	0.0211
	(0.0299)	(0.0177)	(0.0180)
Constant	0.9359***	0.0233**	0.0408***
	(0.0170)	(0.0105)	(0.0133)
Observations	1,135	1,135	1,135

Notes. Standard errors are calculated from 25 replications of a bootstrapping procedure that samples cases (with replacement) clustered at the prosecutor-level and calculated the bunching dummy variables within each sample. The standard errors for the constant terms are not calculated in this way; robust errors clustered at the prosecutor-level are used. 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 (for columns 1-3) and to prosecutors with 10+ cases post-2010. Note, column (2) is a mechanical relationship, hence the missing standard error. 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 (for columns 5-8) and to prosecutors with 10+ cases pre-2010.

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

Table A20. 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 A21. 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 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. 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:

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

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 or if the drug weight is missing and equal to zero otherwise. There is still a decrease in bunching after Alleyne when accounting for missing values.
*** p<0.01, ** p<0.05, * p<0.1

Table A24. Degree of Bunching Post-2010 by Race and District-level Caseload Characteristics

	Pr(280-290g)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After '10 x White (W)	0.0172**	0.0183*	0.0161**	0.0197*	0.0131	0.0219***	0.0113*	0.0137	0.0128
	(0.0082)	(0.0100)	(0.0080)	(0.0102)	(0.0088)	(0.0083)	(0.0061)	(0.0089)	(0.0083)
After '10 x Black or Hispanic (BH)	0.0424***	0.0477***	0.0344***	0.0536***	0.0302***	0.0388***	0.0407***	0.0368***	0.0379***
	(0.0094)	(0.0035)	(0.0028)	(0.0035)	(0.0027)	(0.0029)	(0.0035)	(0.0030)	(0.0033)
After '10 x W x Char.	-0.0147	-0.0088	-0.0055	-0.0122	0.0008	-0.0191*	0.0007	-0.0043	-0.0028
	(0.0098)	(0.0116)	(0.0107)	(0.0115)	(0.0108)	(0.0103)	(0.0104)	(0.0107)	(0.0104)
After '10 x BH x Char.	-0.0187	-0.0222***	0.0027	-0.0363***	0.0124***	-0.0072*	-0.0077*	-0.0011	-0.0031
	(0.0114)	(0.0043)	(0.0043)	(0.0042)	(0.0044)	(0.0042)	(0.0045)	(0.0044)	(0.0044)
Constant	0.0024*	0.0054***	0.0054***	0.0053***	0.0053***	0.0053***	0.0045***	0.0046***	0.0046***
	(0.0014)	(0.0010)	(0.0011)	(0.0010)	(0.0011)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
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 with Appointed Counsel	District Above Med. % Of Cases with Public Defender Counsel
P-value: W = BH	0.0246	0.0057	0.0297	0.0017	0.0609	0.0536	0.0000	0.0139	0.0049
P-value: W+Char. = BH+Char.	0.0000	0.0113	0.0007	0.0872	0.0001	0.0000	0.0191	0.0001	0.0003
Observations	52,731	52,745	52,745	52,745	52,745	52,745	49,851	49,851	49,851

Notes. Robust standard errors in parentheses. "Characteristic" or "Char." represents a dummy variable that is an district or district-by-year characteristic. The specific characteristic of interest is noted in the "Characteristic" row. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. 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, codes 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. Places with different rates of retained, appointed, or public defender defense counsel from 1999-2002 nevertheless have similar bunching at 280g post-2010.

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

Table A25. Relationship between Bunching at 280g and Judge Characteristics

	Pr(280-290g) (1)	Pr(280-290g) (2)	Pr(280-290g) (3)
After 2010	0.0928***	0.0891***	0.1042***
	(0.0093)	(0.0209)	(0.0151)
After 2010 ×White Judge		0.0045	
		(0.0233)	
After 2010 ×Republican Judge			-0.0197
			(0.0191)
Constant	0.0040***	0.0059**	0.0049***
	(0.0007)	(0.0024)	(0.0014)
Observations	8,359	8,359	8,359

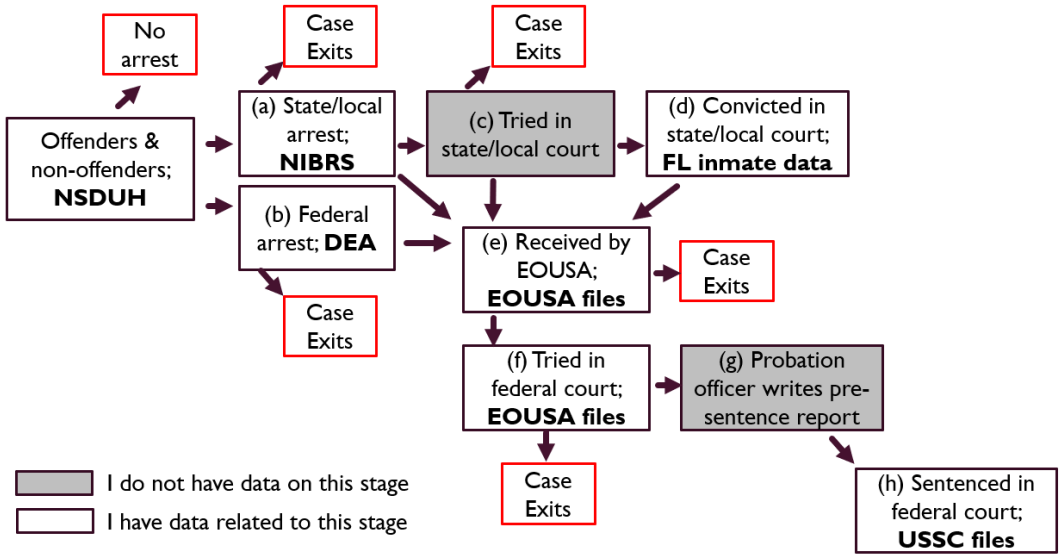
Notes. Standard errors clustered at the judge level in parentheses. The estimates in this table are based on the EOUSA data. I can match judge race and political party to approximately half of the cases in the EOUSA data. For data on judge characteristics, I use the file provided by Cohen and Yang (2019). 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 A26. 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	-0.0129	-0.00857	0.0557			
	(0.0305)	(0.0286)	(0.0412)			
Judge Bunches at 28-29g Post-2010				-0.00207	-0.0144	
				(0.0523)	(0.0329)	
Judge Bunches at 50-60g Pre-2010						0.0175
						(0.0215)
Constant	0.155***	0.143***	0.199***	0.168***	0.108***	0.0723***
	(0.0195)	(0.0185)	(0.0243)	(0.0390)	(0.0250)	(0.0180)
Sample Restriction	0-280g	0-280g, 290-1000g	0-1000g	29-1000g	0-28g, 29-1000g	0-1000g
Observations	769	827	2,710	469	789	1,270

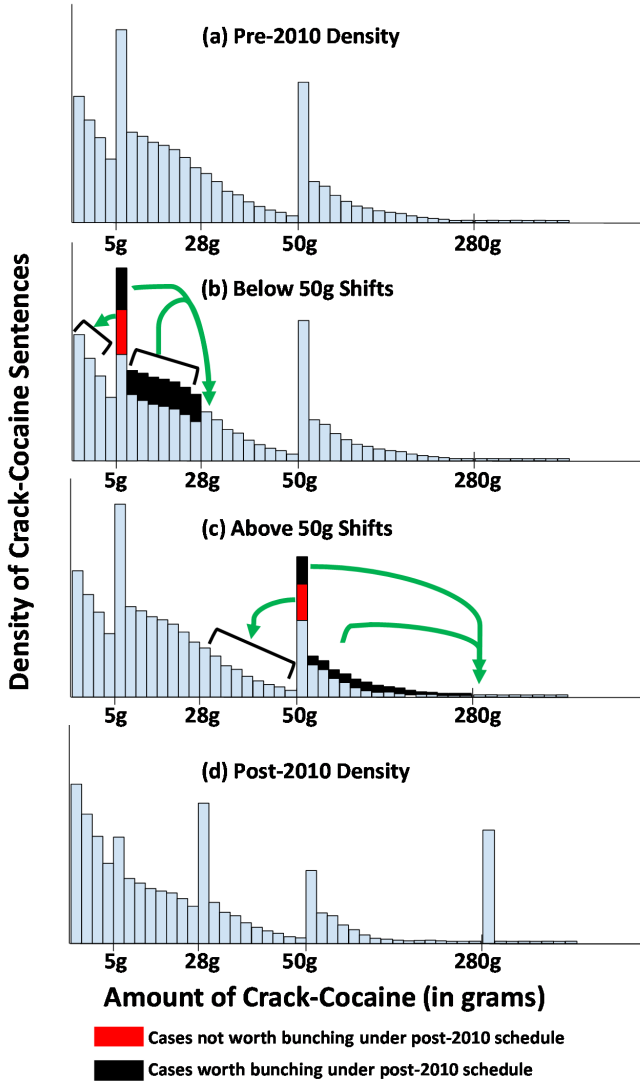
Notes. Standard errors clustered at the judge level in parentheses. The estimates in this table are based on the EOUSA data. See Table A21 for a discussion of the dependent and independent variables in column 1-6. The major difference is that these regressions examine judges classified as "bunching" 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

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



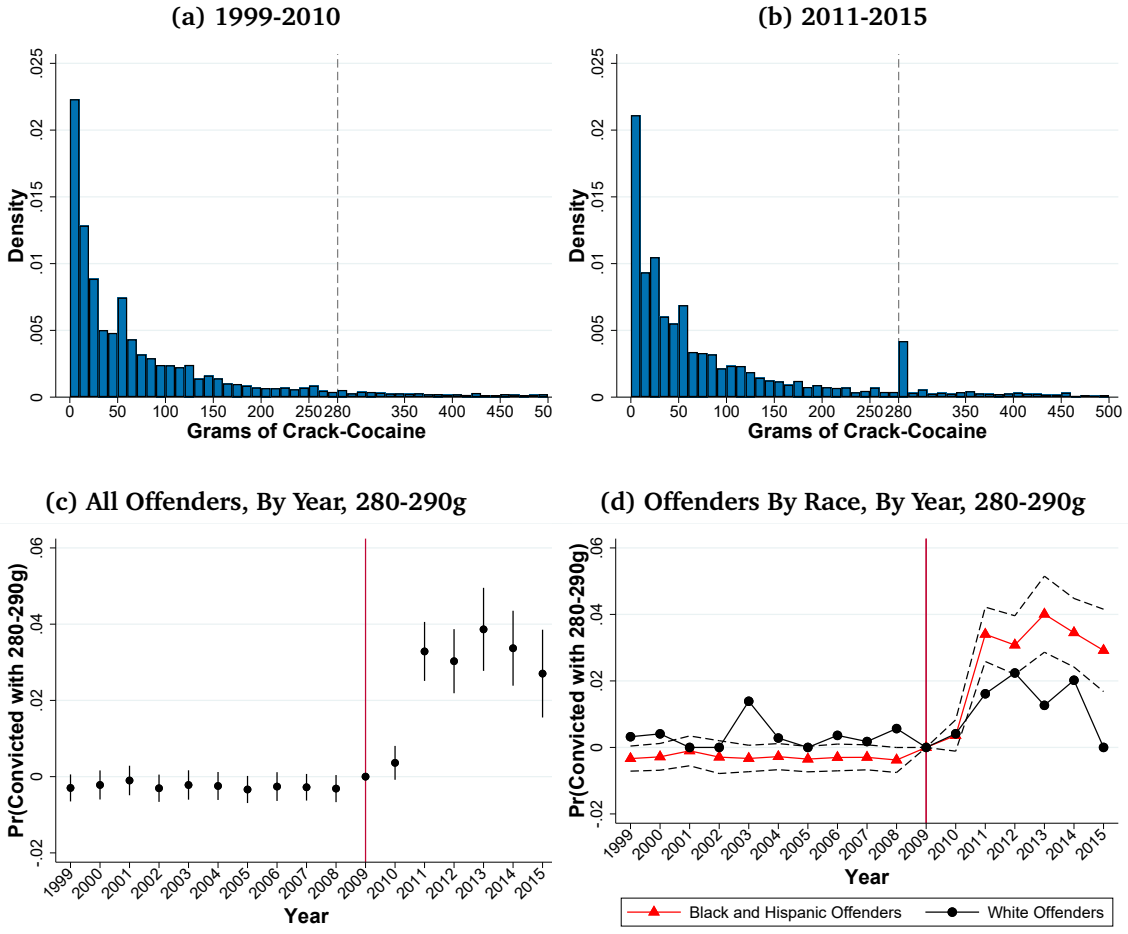
Notes. The figure above details the timeline from arrest to sentencing. Before arrest, the eventual arrestees come from the set of all people, some of whom are innocent and some of whom are guilty. Some individuals from this group are arrested by state/local police or federal police. Of those arrested by state/local police, their case can be dismissed, tried in state/local court, or passed on to federal authorities. Case tried 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 convicted cases can be sent to federal authorities. Individuals arrested by federal police 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 taken to 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.

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



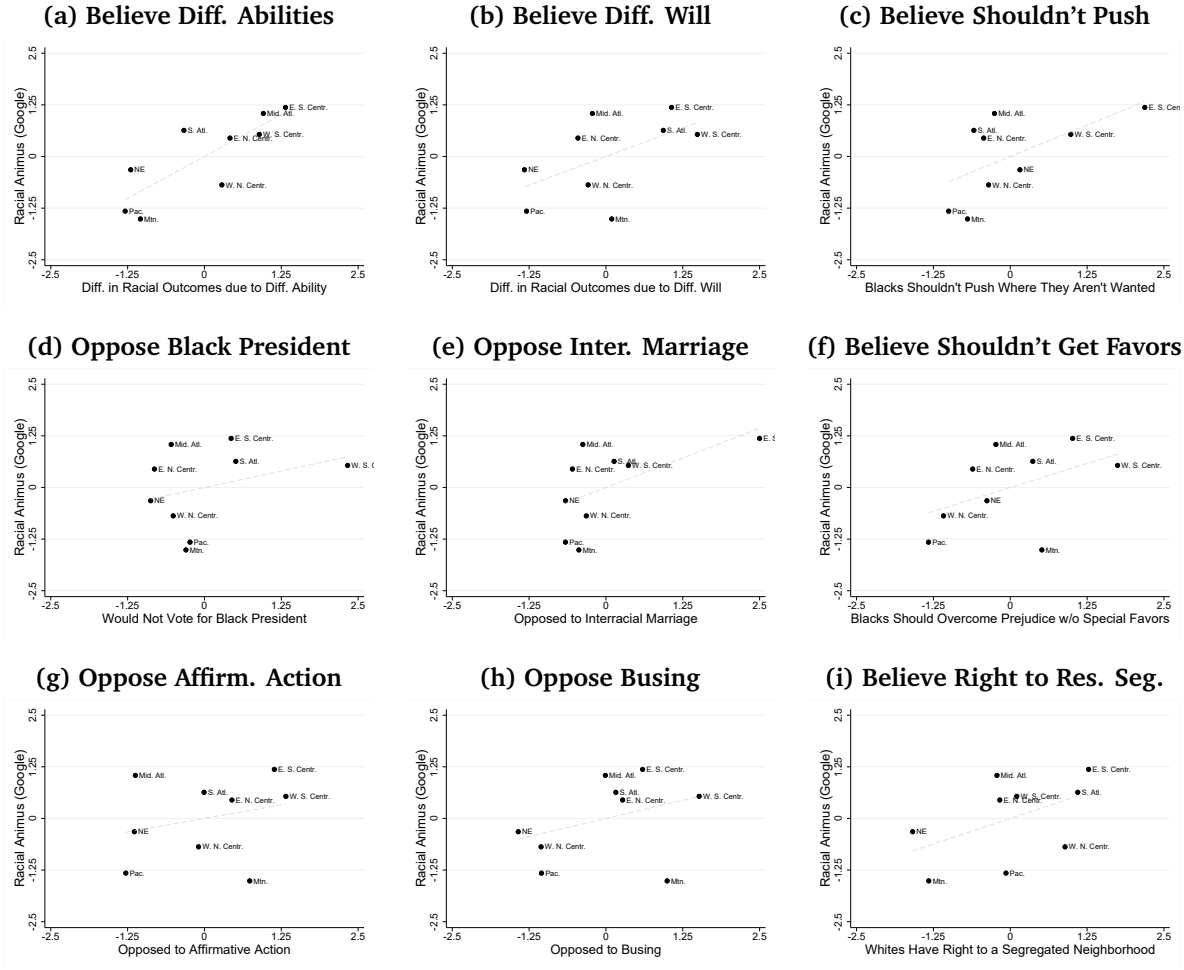
Notes. Panel (a) displays a hypothetical pre-2010 distribution of weights, with bunching at 5g and 50g due to round-number bias and prosecutor discretion. Panel (b) shows how the 0-5g, 5-28g, and 28-50g ranges will change after 2010. Some cases bunched at 5g will not be worth bunching at 28g (depicted 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).

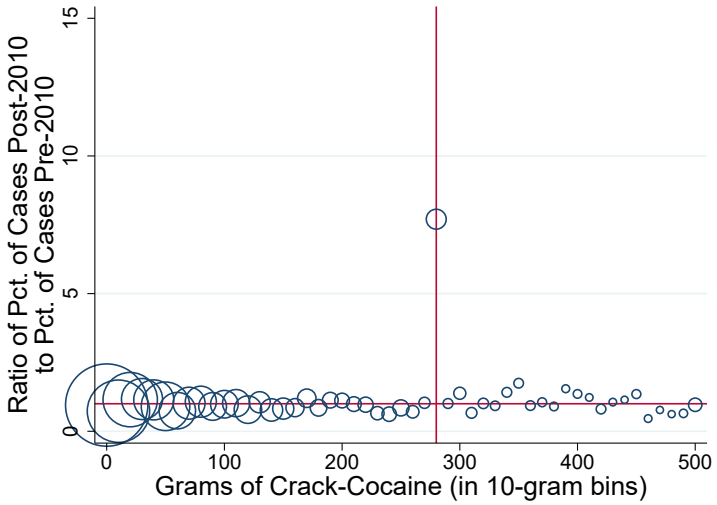
Figure A4. Relationship between Google Trends Racial Animus Measure and GSS Responses from Highly Educated Respondents on Attitudes about Race



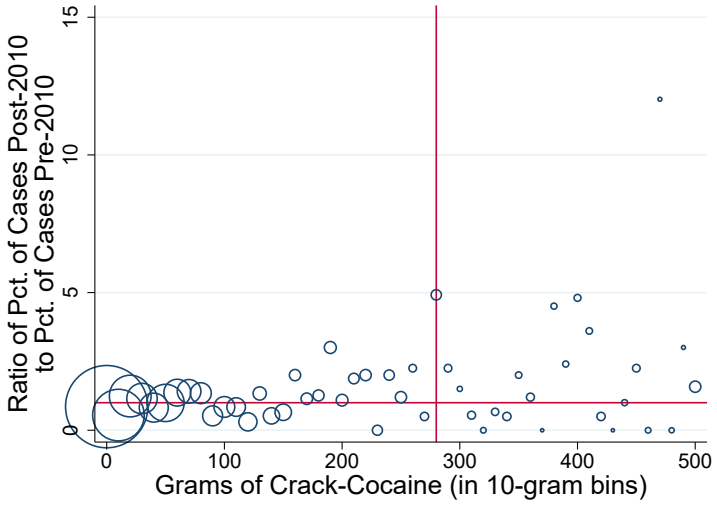
Notes. The figures above plot the relationship between the Google Trends racial animus measure (standardized and centered at zero) and various measures of attitudes about race from the General Social Survey (GSS) from 1972-2018 (not all questions are present in all years; also standardized and centered 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 highly educated people. 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. Bunching Ratio from 0-500g, USSC

(a) Bunching Ratio from 0-500g, Black and Hispanic Offenders



(b) Bunching Ratio from 0-500g, White Offenders

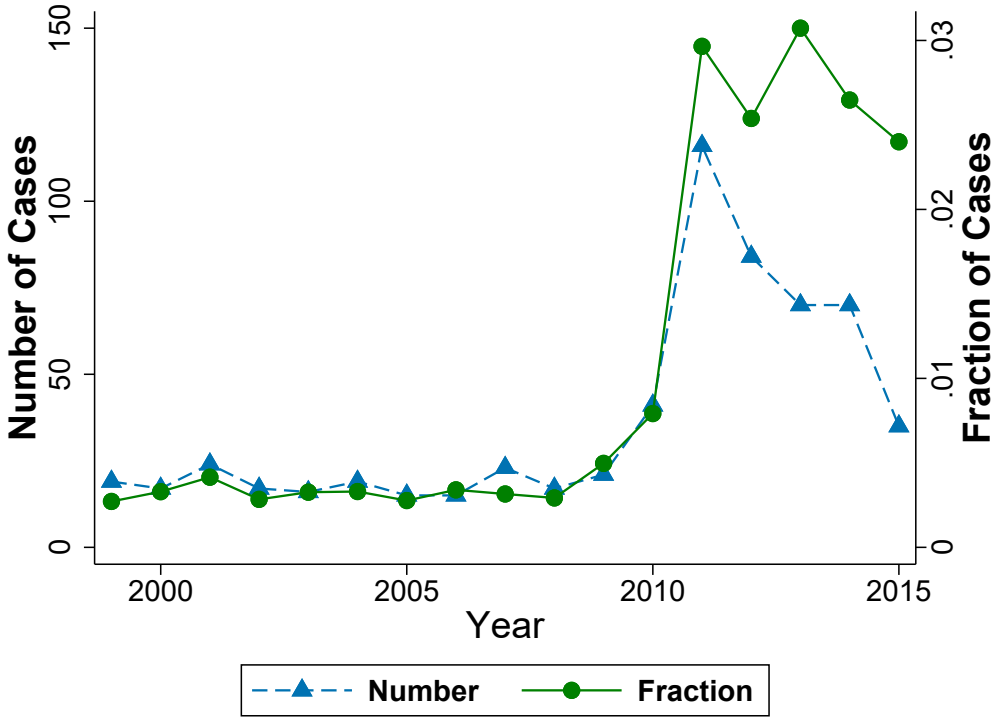


Notes. The figure above plots the bunching ratio for each 10-gram bin from 0-500 grams by race. The bunching ratio for each bin b is defined as follows:

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

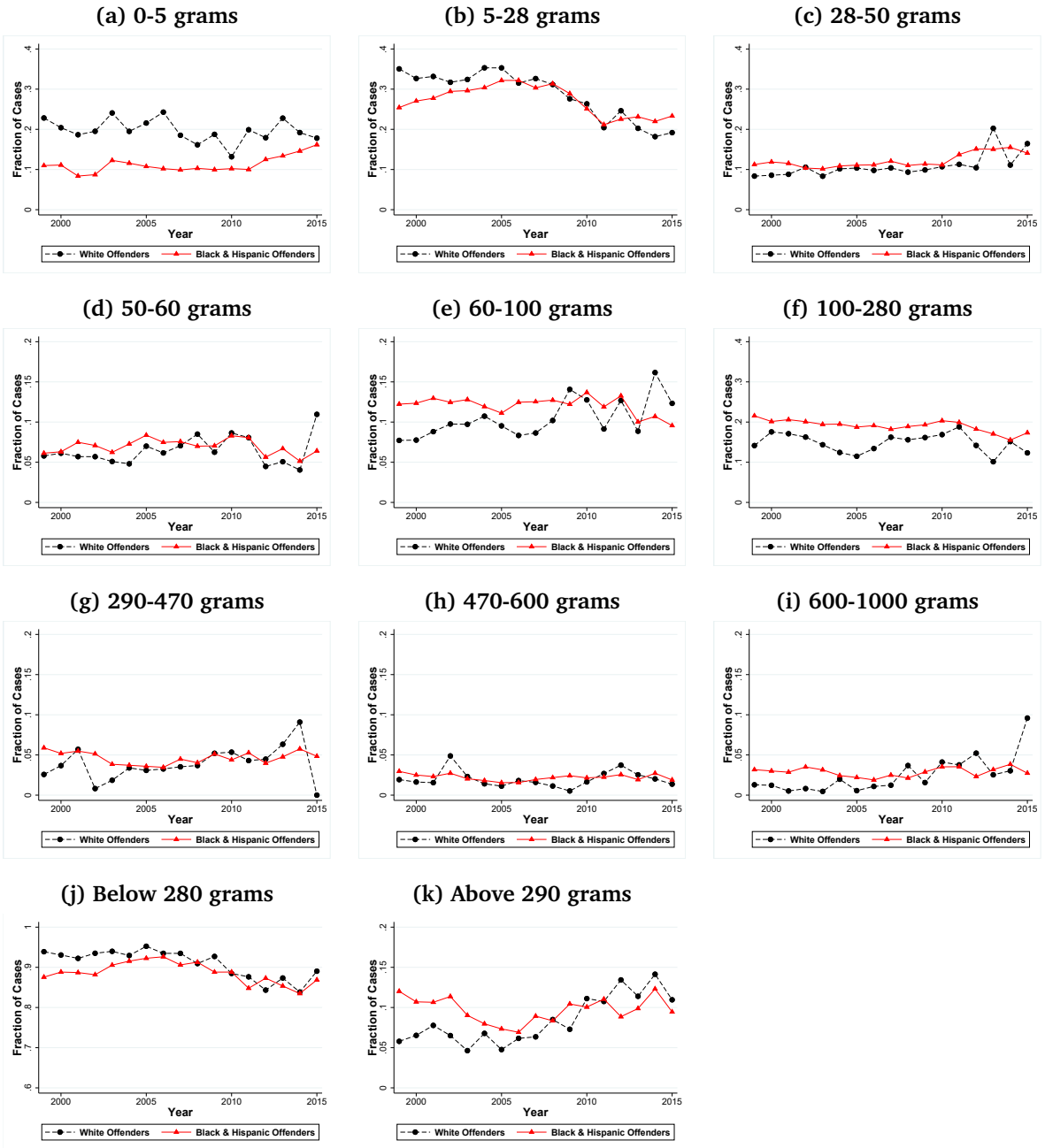
If the distributions are the same pre- and post-2010, the bunching ratio will equal 1 (marked by the horizontal red line). If the ratio is above 1, there is a higher degree of bunching in bin b post-2010. If the ratio is below 1, there is a lower degree of bunching post-2010. The size of the marker for 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).

Figure A6. Number and Share of Offenses with 280-290g Over Time, USSC



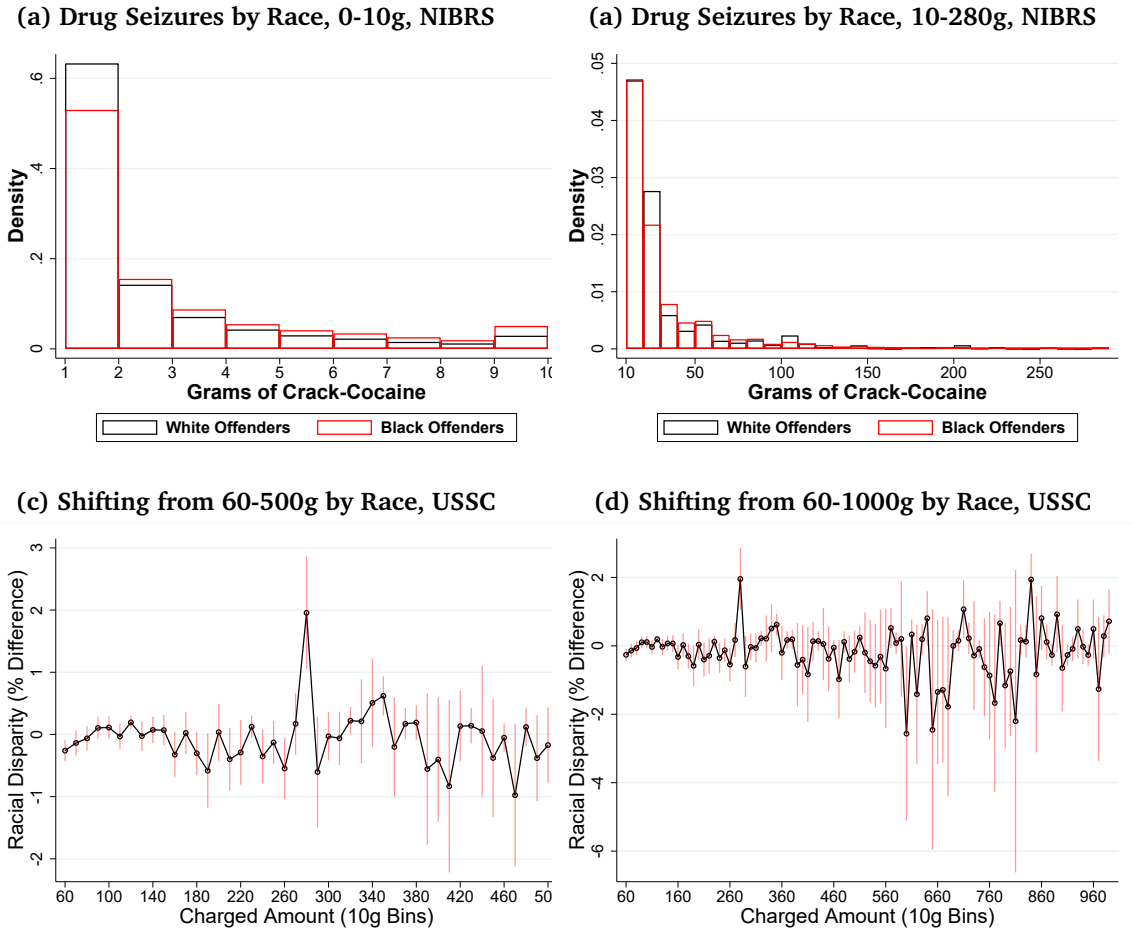
Notes. The figure above plots the total number of offenses with 280-290g over time and the share (or fraction) of cases with 280-290g over time.

Figure A7. 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 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.

Figure A8. Alternative Figures for Conditional Racial Disparity Tests

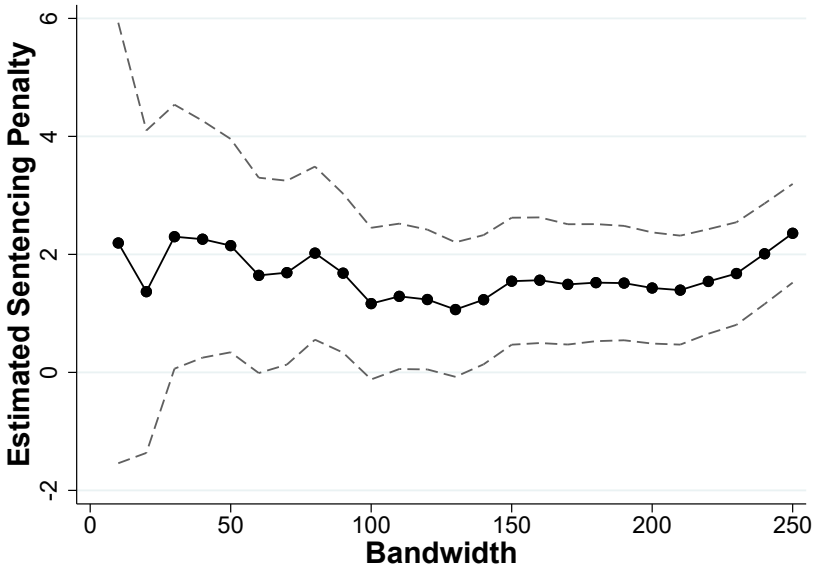


Notes. The figure in panel (a) plots the histograms of crack-cocaine amounts seized for white offenders and for black and Hispanic offenders from 0-10g. The white offenders are slightly over-represented at 1g, but otherwise, the distributions are very similar. The figure in panel (b) plots the histograms 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 for each 10g bin starting at X divided by the share of cases in that 10g bin (to calculate a percent difference).

$$(12) \quad (\text{Charged X-Yg})_{it} = \alpha + \delta^X(\text{After2010}\times\text{BlackOrHispanic})_{it} + \gamma\text{After2010}_{it} + \lambda\text{BlackOrHispanic}_i + \epsilon_{it}$$

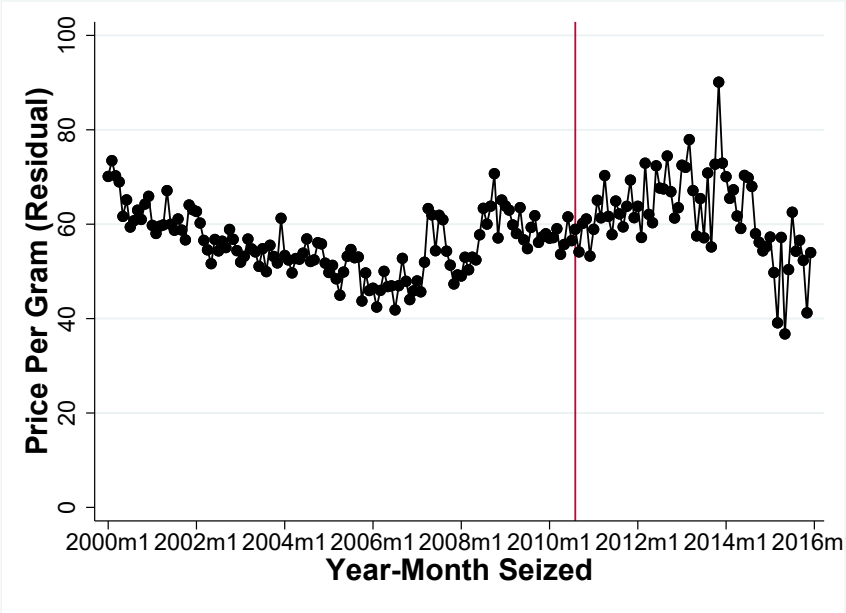
Since estimates are noisier at higher amounts, panel (c) shows the estimates for amounts from 0-500g alone and panel (d) shows the estimates for amounts from 0-1000g.

Figure A9. Sentencing Discontinuity Robust to Multiple Bandwidths



Notes. The figure above plots the sentencing penalty of crossing the 280g mandatory minimum threshold after 2010, as estimated using the RD difference-in-difference model specified in equation (5) of the main text. 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.

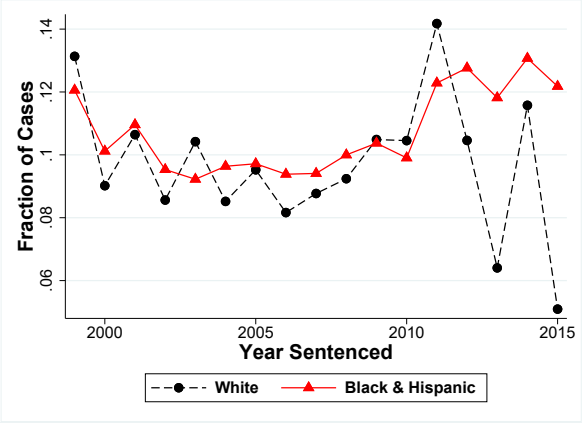
Figure A10. Drug Prices Before and After the Fair Sentencing Act, DEA



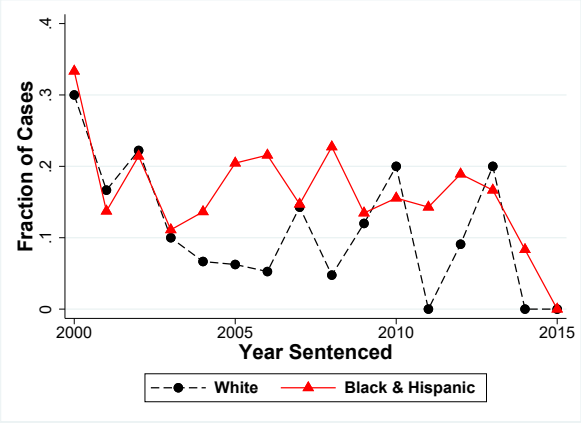
Notes. Panel (a) plots the drug price per gram (conditional on state, drug potency, type of drug, month seized, and a linear trend in year) against the year-month the drugs were seized. Outliers above the 95th percentile ($200 per gram) and below the 5th percentile ($20 per gram) are excluded. The price 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. Panel (b) 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.

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

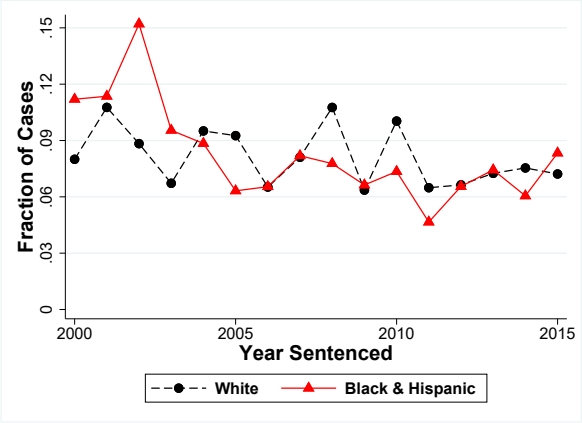
(a) Fraction of Cocaine Cases 200-400g, USSC



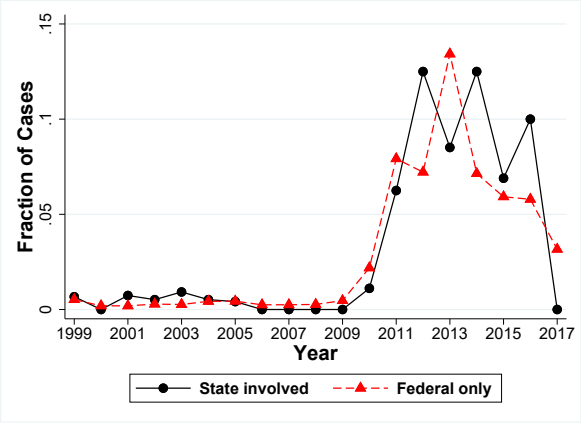
(b) Fraction of Cocaine Cases 200-400g, NC



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

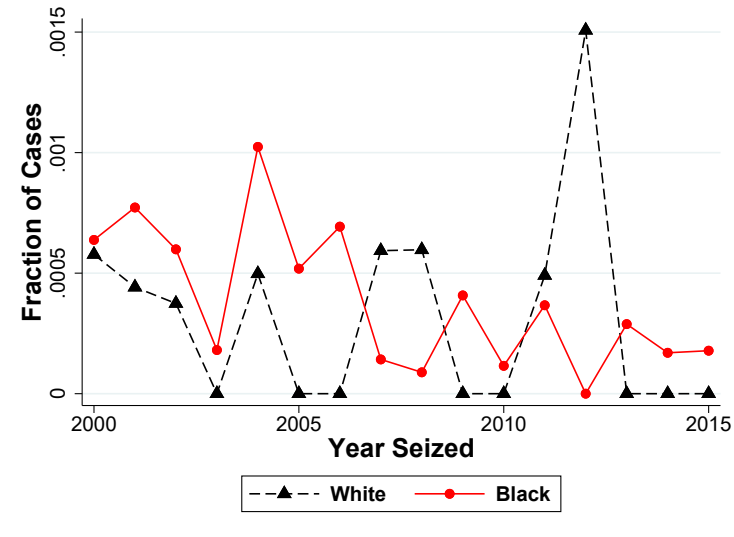


(d) Fraction of Crack-Cocaine Cases in 280-290g, by Type of Source Agency, EOUSA



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 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.

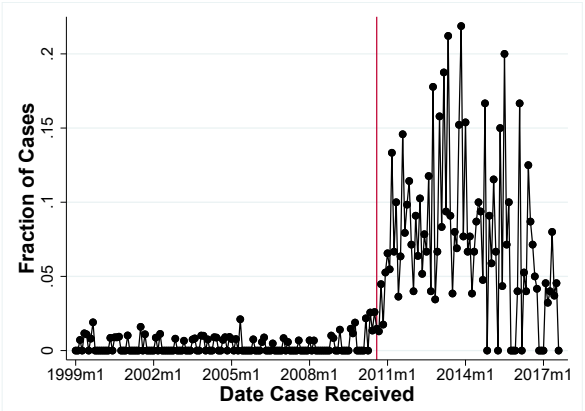
Figure A12. Fraction of Crack-Cocaine Seizures from 280-290g, Full Coverage States, NIBRS



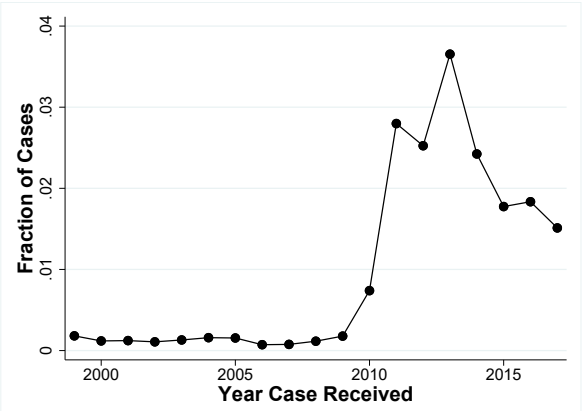
Notes. The figure above plots the fraction of crack-cocaine seizures with 280-290g by race. The sample is limited to states with full coverage (i.e. all agencies in the state participating) starting in 2012 and with 90% coverage or more from at least 2008-2015.

Figure A13. Fraction of Cases with 280-290g Over Time, EOUSA

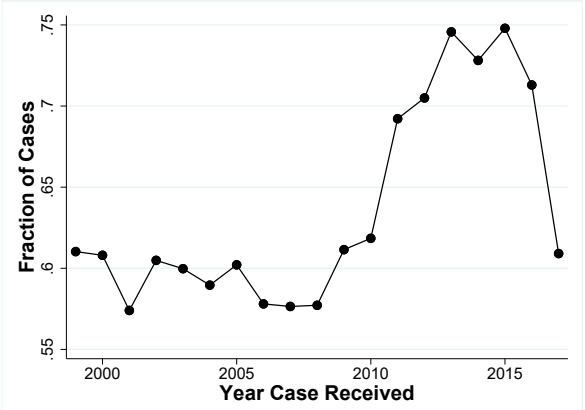
(a) By Month Received by EOUSA



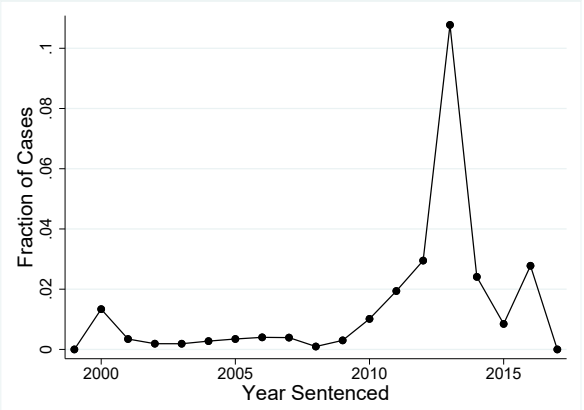
(b) Imputing Missing Weights as (280-290g)=0



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



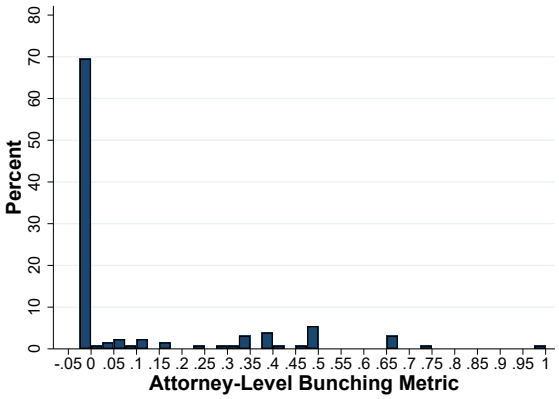
(d) Cases Received Before FSA



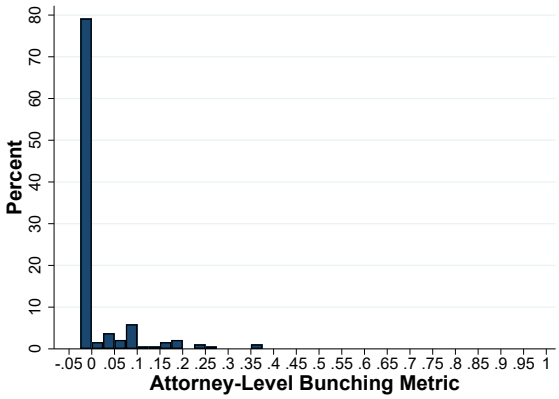
Notes. Panel (a) plots the fraction of cases with 280-290g (excluding cases with missing drug weights) by the month the case was received. The vertical red line indicates the date the Fair Sentencing Act was passed. In panel (b), I re-code the 280-290g dummy variable equal to zero if the drug weight 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. 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.

Figure A14. Histograms of Attorney-level Bunching Metric at 280-290g, EOUSA

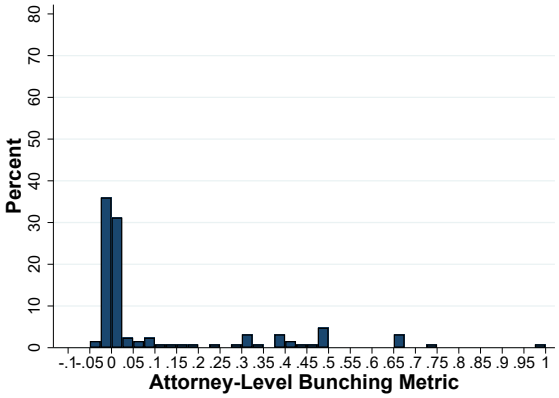
(a) 280-290g, National



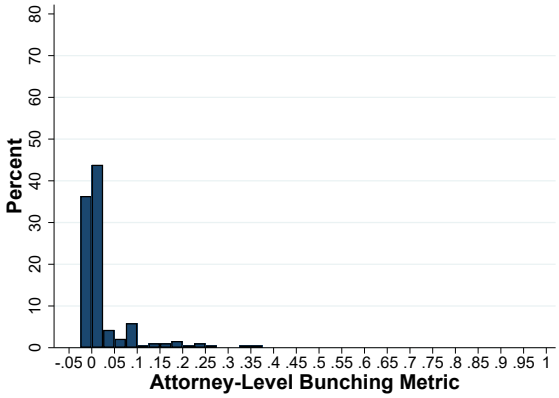
(b) 280-290g, Missing = 0, National



(c) 280-290g, District



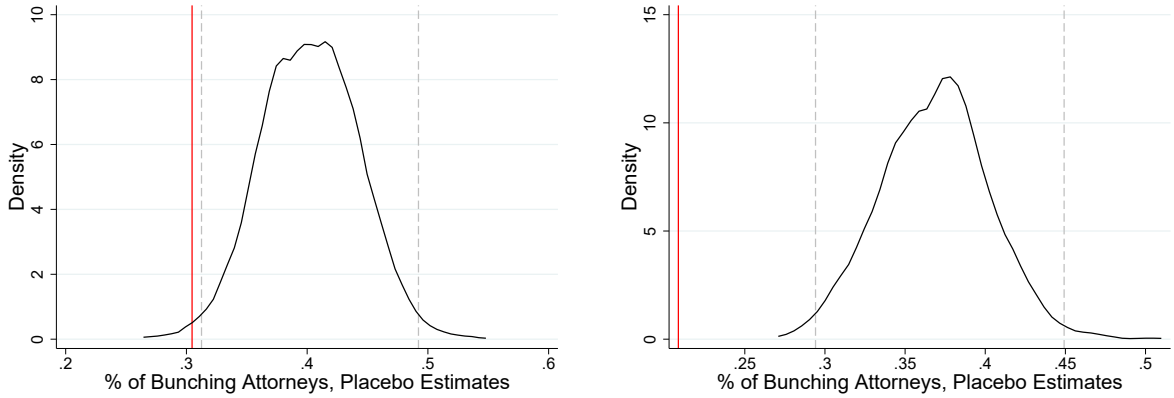
(d) 280-290g, Missing = 0, District



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. 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** 280-290g cases (the highly unrealistic result in Figure A13c) implies that only 70% of attorneys bunch at 280-290g.

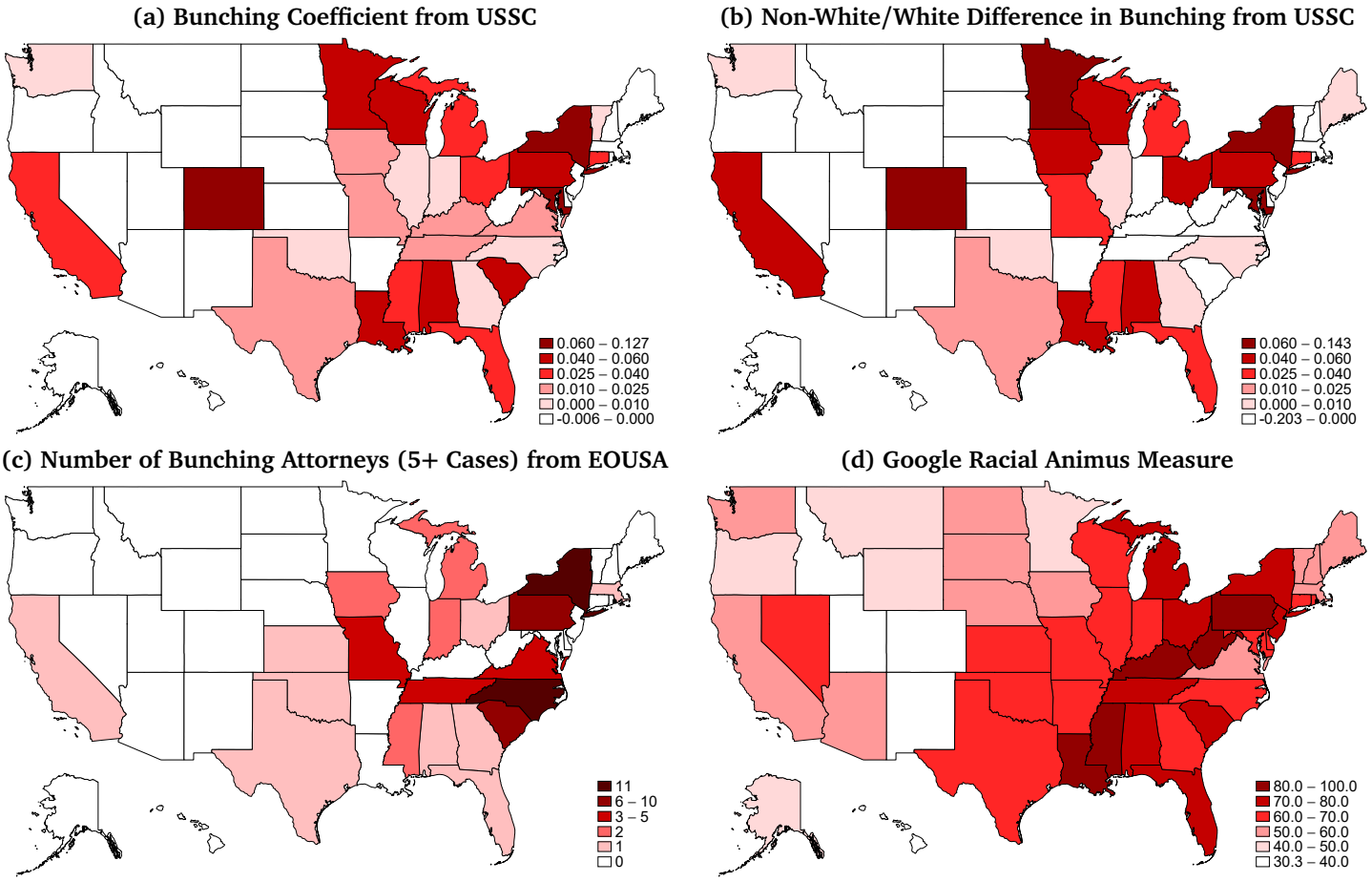
Figure A15. Histograms of Randomized Attorney-level Bunching Metric at 280-290g, EOUSA

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



Notes. 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 from the non-missing data in panel (a) and from the data with missing values imputed in panel (b). The gray dashed lines indicate the 1st and 99th percentiles of the placebo distribution and the red line indicates the fraction of bunching attorneys from the true data.

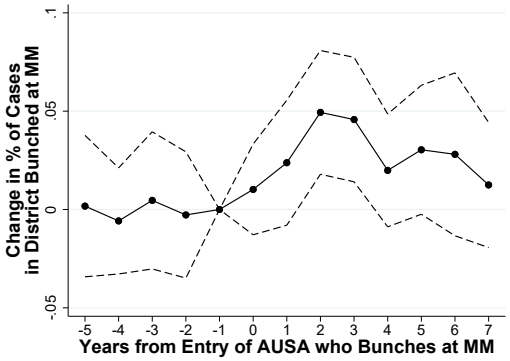
Figure A16. Map of State-level Bunching and State-level Racial Disparity in Bunching



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. Panel (c) plots the number of prosecutors who bunch in each state (among those prosecutors with 5+ drug cases after 2010). 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.

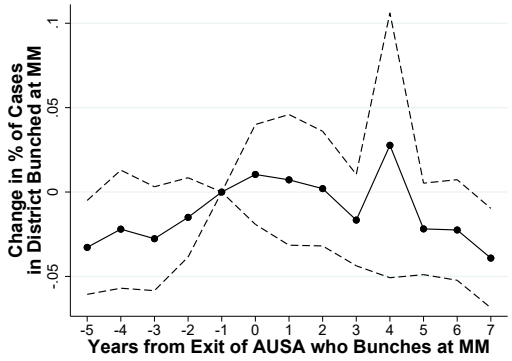
Figure A17. Additional Evidence of Prosecutorial Discretion in Bunching, Alleyne Results and Movers Results, EOUSA

(a) Effect of Entry of a Bunching AUSA



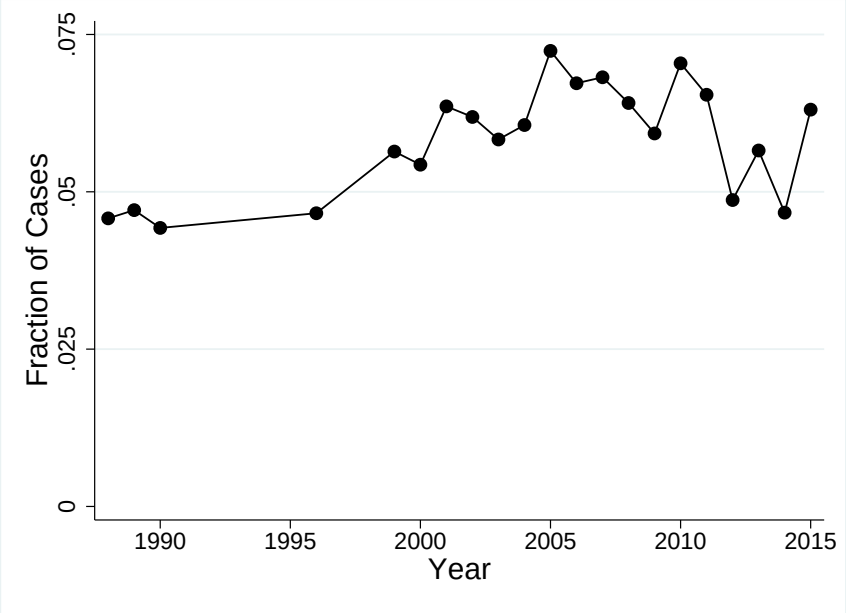
(b) Effect of Entry of a Bunching AUSA, Low-Bunching Districts

(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, all 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.

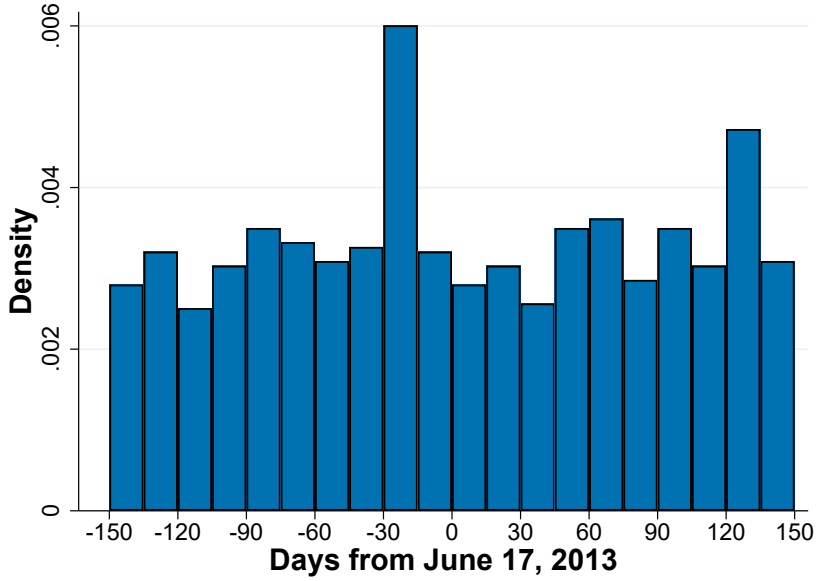
Figure A18. Fraction of Cases in 50-60g by Year, from USSC Sentencing Data



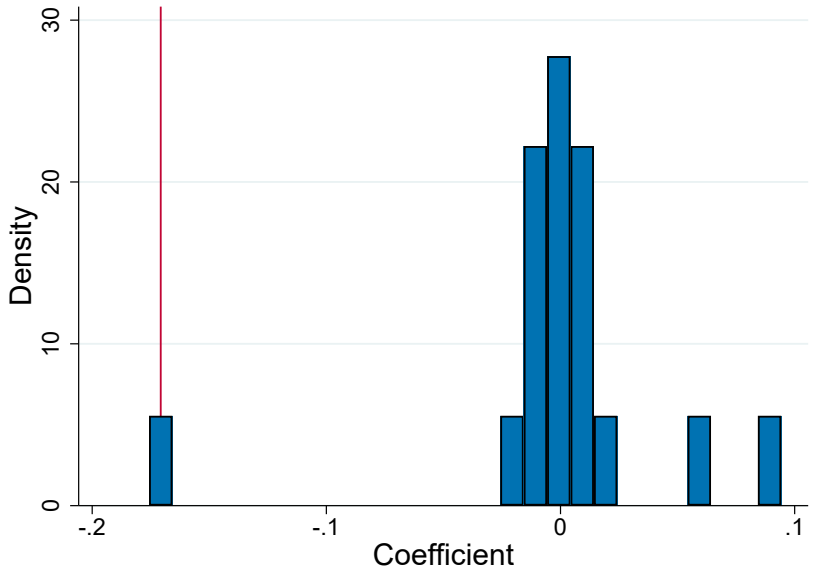
Notes. The figure above plots the fraction of all cocaine (powder and crack) cases with 50-60g by year. The sample is limited to cases with drug weights from 0-1000g. All cocaine cases are used because earlier years (1988-1990) do not distinguish between types of cocaine. This figure indicates that cases 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 A17.

Figure A19. Tests of Validity for Alleyne v. US Result, EOUSA

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

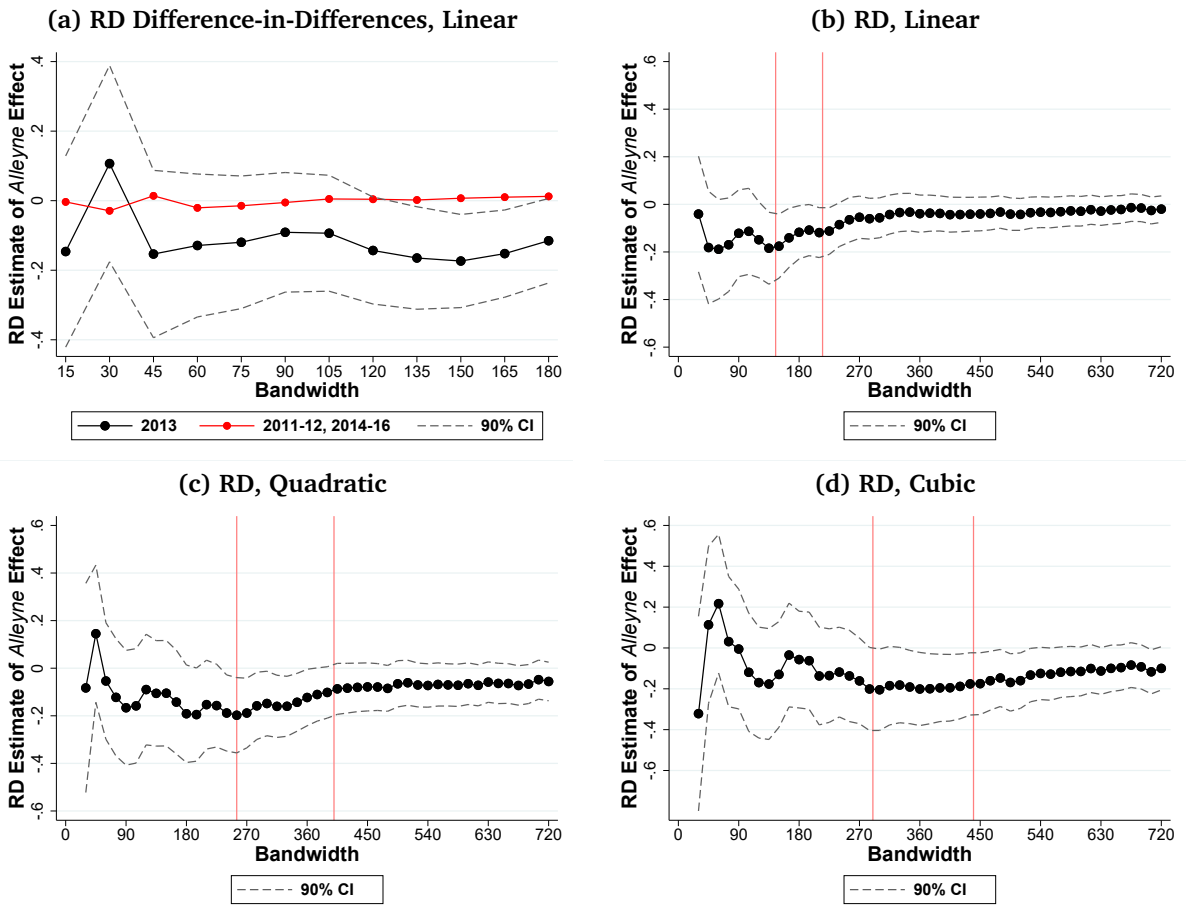


(b) Estimate of Discontinuity Around June 17 in All Years 1999-2016



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.

Figure A20. Robustness of Alleyne v. US Result to Choice of Bandwidth and Polynomial, EOUSA



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. Each panel displays 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 (a) 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 (a) 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 (b)-(d) 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 (Cattaneo et al. 2018). In panel (b), for example, the estimate approaches zero at larger bandwidths–this is to be expected. As we get further from the cutoff, the a 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).

Appendix B. 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 (Best et al. 2015; Fack and Landais 2016; Gelber, Jones, and Sacks 2017; Zaresani 2017; Chen et al. 2018). 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 (Brown 2013; Best et al. 2018; Best and Kleven 2018; Cengiz, Dube, Lindner, and Zipperer 2018). Others compare the distributions using regression analysis on the microdata (Kleven et al. 2011; Behaghel and Blau 2012; Sallee and Slemrod 2012; Chetty, Friedman, and Saez 2013; Dwenger et al. 2016; Goncalves and Mello 2018; and Traxler et al. 2018). 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 employ both direct comparison methods (aggregate/binned analysis and microdata analysis). 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. (2018), I estimate 90% confidence intervals with a bootstrap procedure that samples cases with replacement from the microdata before aggregating to the 10g bin level.[1] 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 B1 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 3.5 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.

In Figure B2a, I plot the difference between the post-2010 and the scaled pre-2010 densities for each 10g bin and add confidence intervals by using 50 bootstrapped samples from the microdata. In addition, I also display a table of the statistical results for the binned missing mass analysis in Figure B2b. When this difference is below zero, it means the bin contains relatively fewer cases after 2010 and when the difference is above zero, it means the bin contains more cases after 2010.

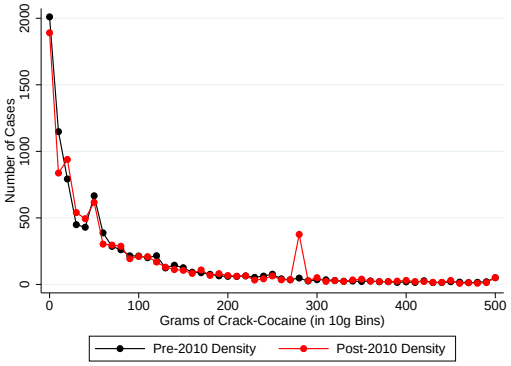
The figure shows an increase of about 340 cases in the 280-290g bin post-2010, a net increase in cases above 280g, and a net decrease below 280g. Summing the changes in bins above 280g, I find a net increase in that section of the distribution after 2010. The point estimate on the net change is noisy, but even summing the lower bound of the confidence interval for all bins above 280g can only account for about 46% of the increase in the 280-290g bin. On the other hand, the net change below 280g can account for 120% of the increase in the 280-290g bin. Again, this point estimate is noisy. In fact, summing the upper confidence interval for all bins below 280g implies

[1] 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.

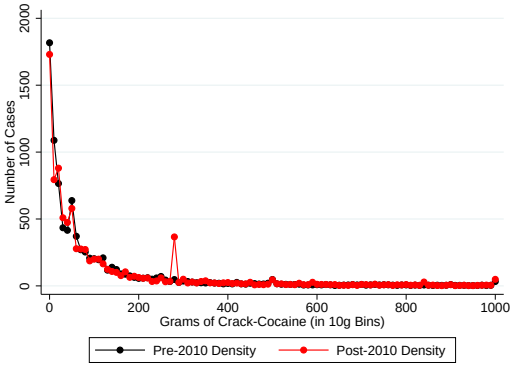
a net increase in that section of the distribution. The key takeaway is that changes in the distribution below 280g can account for the excess mass at 280g, whereas changes in the distribution above 280g cannot. In other words, an offender charged with 280-290g post-2010 would likely have been charged with less than 280g had they been sentenced prior to 2010. Table B1 displays the results from similar binned analyses using the NIBRS data, DEA data, and EOUSA data.

Figure B1. Scaled Pre-2010 Distribution of Recorded Weights vs. Post-2010 Distribution

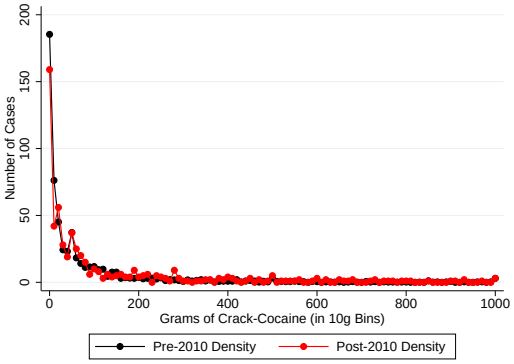
(a) All Offenders



(b) White Offenders

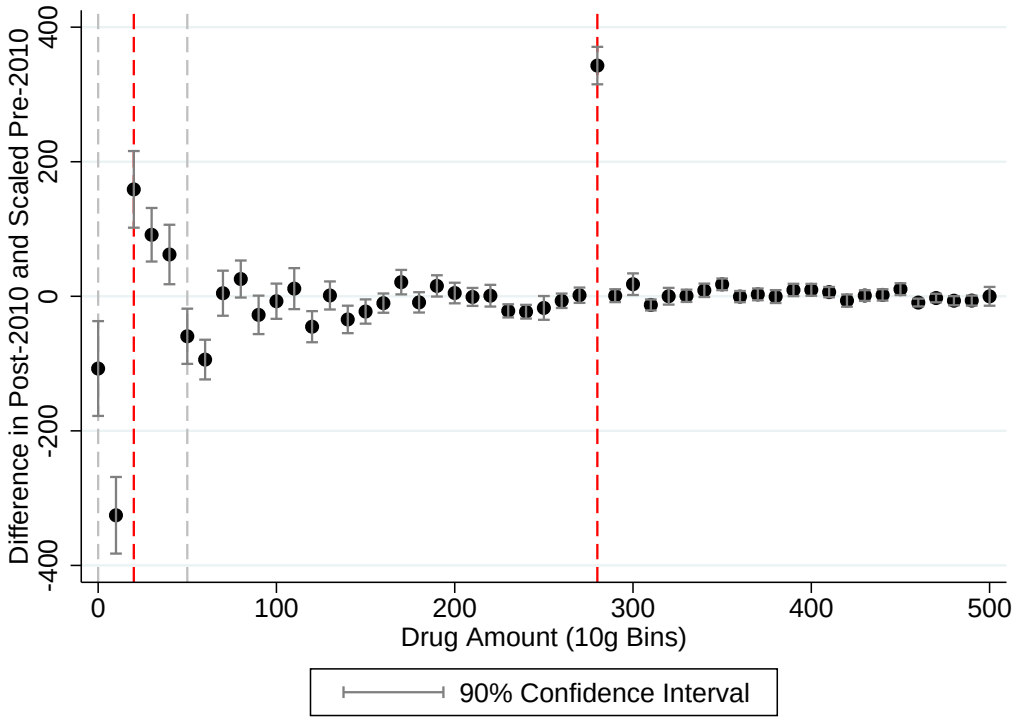


(c) Black and Hispanic Offenders



Notes. Figure B1a 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 1000g. Figures B1b and B1c do the same but restrict the sample to white offenders or black and Hispanic offenders, respectively.

Figure B2. Post-2010 Density Minus Scaled Pre-2010 Density
(a) Difference between Post-2010 and Pre-2010 Distribution of Drug Amounts



(b) Fraction of Bunching Accounted for by Different Ranges

Range	Net Difference	90% CI	% Bunching at 280g
0-20g	-435.56	(-558.17, -312.94)	128.67%
20-50g	293.49	(158.10, 428.88)	-86.70%
50-60g	-52.63	(-101.67, -3.59)	15.55%
60-100g	-65.00	(-184.58, 54.59)	19.20%
100-280g	-122.43	(-414.25, 169.39)	36.17%
0-280g	-382.13	(-1100.58, 336.32)	112.89%
290-500g	43.44	(-146.74, 233.62)	-12.83%

Notes. The figure above plots the difference between the post-2010 density and the scaled density of drug quantities in pre-2010 for each 10-gram bin. Confidence intervals are calculated by bootstrapping as discussed in the text. The red dashed lines correspond to the post-2010 mandatory minimum bins (28g and 280g) and the gray dashed lines correspond to the pre-2010 mandatory minimum bins (5g and 50g). Summing the changes in bins above 280g, I find a net increase in that section of the distribution after 2010. The point estimate on the net change is noisy, but even summing the lower bound of the confidence interval for all bins above 280g can only account for about 46% of the increase in the 280-290g bin. On the other hand, the net change below 280g can account for 120% of the increase in the 280-290g bin. Even the changes from 50-280g can account for 85% of the increase in the 280-290g bin. Panel B displays statistical results for relevant drug amount ranges.

Table B1. All Bunching Results using Aggregated/Binned Comparison with Bootstrapped SEs

	Pr(280-290g Crack-Cocaine Recorded)					
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0347***		-0.0002***		-0.0006***	0.0771***
	(0.0020)		(0.0001)		(0.0002)	(0.0054)
After 2010 x White		0.0126**		-0.00002		
		(0.0062)		(0.0001)		
After 2010 x Non-White		0.0359***		-0.0003***		
		(0.0023)		(0.0001)		
Constant	-0.0003***	-0.0001***	0.000002	0.0000002***	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	57,101	52,940	203,700	203,700	100,306	24,493

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

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 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.

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

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 (8) (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 Figures B3 and B4.

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 $\beta = 0.0352$ and $SE_\beta = 0.0169$.

Next, I estimate:

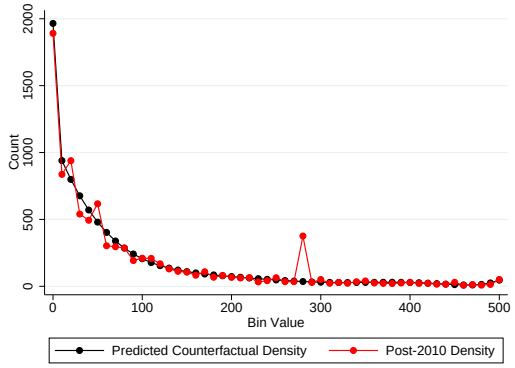
$$(\%\text{ in Post2010 - }\%\text{ in Counterfactual})_{br} = \alpha + \beta Bin280_b + \gamma NonWhite_r + \delta Bin280_b \times NonWhite_r + \epsilon_b$$

Using the Saez (2010) and Chetty et al. (2011) method, I estimate $\delta = 0.0237$ and $SE_\delta = 0.0119$. Using

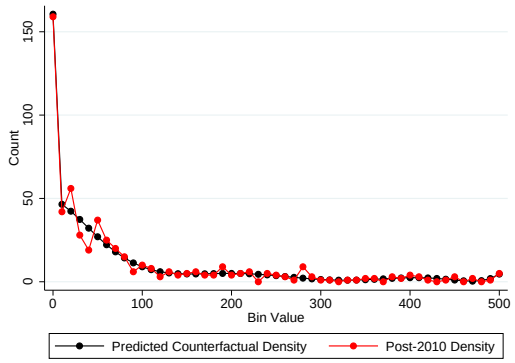
the difference-in-bunching method, I estimate $\delta = 0.0216$ and $SE_\delta = 0.0109$. In all analyses, I detect substantial bunching after 2010 and disproportionate bunching after 2010 for black and Hispanic offenders.

Figure B3. Predicted Counterfactual Density and Post-2010 Density

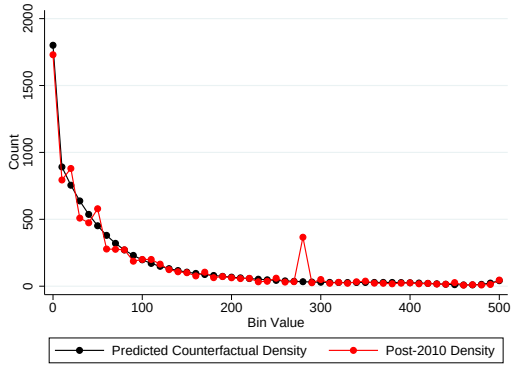
(a) All Offenders



(b) White Offenders, Saez (2010) Method



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



Notes. In panel (a), I plot a predicted counterfactual density of drug quantities (in black) and the actual density of drug quantities post-2010 (in red). In panels (b) and (c), I 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.

References

Behaghel, Luc, and David M Blau. 2012. "Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age." *American Economic Journal: Economic Policy* 4 (4): 41–67.

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy* 123 (6): 1311–55.

Best, Michael and Henrik Jacobsen Kleven. 2018. "Housing Market Responses to Transaction Taxes: Evidence from Notches and Stimulus in the UK." Working paper.

Best, Michael, James Cloyne, Ethan Ilzetzki, and Henrik Jacobsen Kleven. 2018. "Interest Rates, Debt and Intertemporal Allocation: Evidence from Notched Mortgage Contracts in the United Kingdom." Working paper.

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

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2018. "The Effect of Minimum Wages on Low-Wage Jobs: Evidence from the United States Using a Bunching Estimator." Working paper.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suarez Serrato, and Daniel Yi Xu. 2018. "Notching R&D Investment with Corporate Income Tax Cuts in China." Working paper.

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

Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103 (7): 2683–2721.

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 2016. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." *American Economic Journal: Economic Policy* 8 (3): 203–32.

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

Gelber, Alexander M., Damon Jones, and Daniel W. Sacks. 2013. "Estimating Earnings Adjustment Frictions: Method and Evidence from the Social Security Earnings Test." NBER working paper.

Goncalves, Felipe and Steven Mello. 2018. "A Few Bad Apples? Racial Bias in Policing." Working paper.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Soren Pedersen, and Emmanuel Saez. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." 2011. *Econometrica* 79 (3): 651–692.

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

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

Sallee, James M., and Joel Slemrod. 2012. "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." *Journal of Public Economics* 96 (11): 981-999.

Traxler, Christian, Franz G. Westermaier, and Ansgar Wohlschlegel. 2018. "Bunching on the Autobahn? Speeding Responses to a 'Notched' Penalty Scheme." *Journal of Public Economics* 157: 78–94.

Zaresani, Arezou. 2016. "Adjustment Costs and Incentives to Work: Evidence from a Disability Insurance Program." University of Calgary working paper.

Appendix C. Supplementary Materials for Model of Prosecutor Objectives

I. Prosecutor Responses to Changing Mandatory Minimum Thresholds

The model from Section V.E also has implications about how the optimal choice in period $t = 0$ relates to the optimal choice in period $t = 1$. 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 1 shows these possible rational choices of a^{1*} for a given a^{0*} and ranges of s.

Table 1. 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.

II. Prosecutors' Signal Extraction Problem

The racial disparity in bunching at 280g after 2010 could be due to statistical discrimination. 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:

$$d_r \sim N(\bar{d}_r, \sigma_d^2)$$
$$\bar{d}_{bh} > \bar{d}_w$$

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 $l^*(s,r,x)$, 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 $l^*(s,r,x=x_1) > l^*(s,r,x \neq x_1)$.

Appendix D. Data Appendix

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. 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. According to USSC, if the information in these documents conflicts, the JC takes precedent.

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. The values for WGT{X} are converted from variables DRGAM{X} and UNIT{X}. Information on drug amount and drug unit is taken from the Statement of Reasons (SOR), if present. If not present in the SOR, the information is taken from the PSR. According to the USSC, if the information in these documents conflicts, the SOR takes precedent.

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. In fact, 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 guideline manual used for sentencing guidelines calculations. This information is taken from the PSR.

District. The district used for analyses is derived from the variables 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. If both documents are available, and the information conflicts, the JC takes precedent.

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. Key variables from the data are outlined below:

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.

In robustness tests, I use similar data from North Carolina. It also contains an offense string that provides information about drug type and quantity. However, the string does not always specific the type of drug. 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.[1] 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.

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: icpsr.umich.edu/icpsrweb/NACJD/series/128. 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.

[1] **Wright, Ronald.** 2006. "Federal or State? Sorting as a Sentencing Choice." *Criminal Justice* 21 (2): 16-21.

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, in which race is self-reported.

For the primary analyses of the NIBRS data, I limit the sample to a balanced panel of agencies. For robustness checks, I limit to stats that have had full agency coverage in NIBRS since 2012 and over 90% coverage since 1998.

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. 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."

EOUSA Case Management Files

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.
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, 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 highly 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.

Staff ID/Assignment. The EOUSA data also contains an ID variable for the lead attorney assigned to the case. This ID is tied to the district. In other words, two attorneys can have the same numeric ID as long as they are in different districts. Also, this ID will not follow an attorney from one district to another.

Initials. Since the EOUSA numeric 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. In practice, this initials-based ID appears to accurately identify attorneys who switch districts. First, attorneys who move from one district to another continue to bunch at 280g in the new district. Second, when an attorney moves into a new district, other attorneys in that district start to bunch more at 280g. Third, attorneys who I identify as "moved" are often disconnected from their old district in the data and connected to their new district. If the initials-based ID were totally random, we should not expect to see these three patterns.

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:

https://test.openicpsr.org/openicpsr/project/114590/version/V1/view?path=/openicpsr/114590/fcr:versions/V1/Data_2015_0150/Public-Use-Data&type=folder