Racial and Ethnic Sentencing Differentials in the Federal Criminal Justice System*

Brendon McConnell Imran Rasul

January 2018

*We gratefully acknowledge financial support from the Dr. Theo and Friedl Schoeller Research Center for Business and Society, and the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS (grant number RES-544-28-5001). All errors remain our own. Author affiliations and contacts: McConnell (Southampton, brendon.mcconnell@gmail.com); Rasul (University College London, i.rasul@ucl.ac.uk).

1 Introduction

Minority groups are far more likely to come into contact with the Federal criminal justice system (CJS) than Whites. A voluminous body of research further shows sentencing *outcomes* also vary tremendously for such groups. The research challenge lies in establishing whether these sentencing differentials are driven by unobserved heterogeneity correlated to defendant race/ethnicity, or whether they reflect discrimination. We add to the debate by examining the robustness of racial/ethnic sentencing gaps, by gender, when allowing for selection on unobservables (SoU). We use the *Monitoring of Federal Criminal Sentences* (MFCS) data, covering Federal Court cases. Three features make it amenable to examining the robustness of sentencing differentials to SoU.[1]

First, the data covers over 270,000 Federal criminal cases up for sentencing between 1998 and 2003. This allows Black-White and Hispanic-White differentials to be studied. It covers cases from all 90 mainland US Districts, defendants of all ages, genders, and all types of criminal offense. Second, the data contains rich information for each criminal case: defendant demographics include their age, highest education level, marital status and number of dependents. Legal controls include the type of defense counsel and the Federal court district of sentencing. Offense details allow us to classify offenses into 31 types. Third, sentencing guidelines are in place in the Federal CJS in our study period. Such guidelines provide for *determinate* sentencing, mapping combinations of the offense severity and the defendant's criminal history into a specific sentencing range, as shown in Table A1. The *MCFS* data records which of the 258 guideline cells is recommended to the judge pre-sentencing. This effectively proxies case-specific factors the prosecution and legal counsel deem judges should factor into sentencing.[2]

The ability to control for this heterogeneity helps explain much of the predictable variation in sentencing outcomes. Along with the large sample sizes, this enables us to effectively apply SoU methods. We do so to shed light on the plausibility of whether racial/ethnic sentencing differentials, by gender, could be entirely due to unobservable heterogeneity.[3]

[1] The MFCS data links four administrative data sets covering the arrest/offense stage before an individual enters the Federal CJS, and all subsequent stages through to sentencing [USSC MCFS 1999-2003].

[2] For example, an individual who commits a robbery is allocated 20 points. If a gun is involved, 5 more points are awarded. If the individual had been a minimal participant in the robbery, 4 points are deducted. If the individual is found to be in obstruction of justice, 2 more points are added. In this case the final offense severity score would be 23 points. There are six criminal history categories. Criminal history points are based on each prior sentence of imprisonment (and vary with the length of that earlier imprisonment), whether the offenses was committed while under parole/release etc. Suppose the individual in the example above was assessed to have 7 criminal history points. The sentencing guidelines would then stipulate they should be sentenced in the range of 70-87 months, as shown in Table A1.

[3] The consideration of SoU layers onto the analysis of sentencing differentials by race, ethnicity and gender in the FCJS in Sorensen *et al.* [2012].

2 The Federal Criminal Justice System

Criminal cases are filed in Federal court if an individual is prosecuted by a Federal agency or breaks a Federal law. These cases tend to be more serious than those in State courts, and hence their sentencing severity is harsher. Federal Judges are nominated by the President, confirmed by Congress, and life appointees. They are among the most senior judges in the country. Jury trials in Federal courts occur only if a defendant pleads not guilty. This is rare: 96% of defendants plead guilty before they reach trial. By pleading guilty, the individual is convicted and only their sentence remains to be determined. Guilty pleas can be taken into account at sentencing.

Sentencing guidelines were introduced in the Sentencing Reform Act of 1984 by the US Sentencing Commission (USSC). The goal was to alleviate sentencing disparities because guidelines providing for *determinate* sentencing, limiting the discretion judges had over penalties imposed at the sentencing stage. The sentencing guidelines are based on: (i) the severity of the offense; (ii) the defendant's criminal history. Table A1 shows the full set of guideline cells, mapping each possible combination of offense severity (1 to 43) and criminal history (scores 1 to 13, grouped into 6 bins) into a sentencing range. Hence there are 43 x 6 = 258 guideline cells. The lowest of these cells allow for non-custodial sentences (in Zone A); the most severe cells permit only life sentences. The sentencing guidelines still provide judges some discretion, allowing them to *downwards depart* from the recommended guideline cell (so moving in a Northerly direction in Table A1).[4,5]

3 Results

To estimate Hispanic-White and Black-White sentencing differentials, we use two variables available at the sentencing stage in the *MFCS* data. Defendants are classed as either Hispanic or non-Hispanic. A separate race code identifies defendants as white-race, black-race, other-race. Whites are coded as white-race and non-Hispanic; Blacks as black-race and non-Hispanic; Hispanics as white- or black-race and Hispanic. This implies that among male defendants, 31% are White, 26% are Black and 43% are Hispanic. Among female defendants, 41% are White, 31% are

[4] A judge can downwards depart if they find mitigating circumstances not adequately taken into consideration by the USSC in formulating the guidelines. These can include diminished capacity or rehabilitation after the offense but prior to sentencing, family responsibilities or prior good works. Downward departures may also be warranted "*[i]f reliable information indicates that the defendant's criminal history category substantially over-represents the seriousness of the defendant's criminal history or the likelihood that the defendant will commit other crimes.*" Judges are required to provide written explanations for the specific reason(s) for downward departing.

[5] The guideline cells were in operation from 1987 until 2005. The Supreme Court's 2005 decision in US v. Booker found the guidelines violated the Sixth Amendment right to trial by jury. The guidelines are now only considered advisory. Much of the sentencing boom in the State CJS has been attributed to moves towards determinate sentencing, which is argued to more negatively impact outcomes for Blacks [Neal and Rick 2015].

Black and 28% are Hispanic. Table A2 show sample descriptives. These are quite similar across genders. Whites tend to be older, Blacks more likely to be single, and Hispanics with the lowest levels of education. Finally, criminal history is highest among Black defendants, although Black men and Hispanic women have the highest rates offense severity.

Table 1 shows ethnic sentencing differentials by ethnicity and gender. Focusing first on men, Columns 1 to 3 show the unconditional sentencing differentials are large. Blacks are 4.7pp less likely to be downward departed, 12pp more likely to receive some prison sentence, and their prison sentences are 40 months longer than for Whites. Unconditional Hispanic-White sentencing differentials are equally stark in terms of the likelihood of any prison sentence and being downwards departed, although sentence lengths are no different to Whites.

We next examine whether these differentials are robust to conditioning on the observables described earlier. The *MCFS* data allows us to condition on the full set of guideline cells, effectively proxying all case-specific factors that prosecutors and legal counsel deem judges should factor into their sentencing decision (such as whether a gun was used in the crime, the quality of drugs involved in drug offenses etc.). There are large changes in Black-White and Hispanic-White sentencing gaps as we condition on observables. This is as expected given defendants differ in observables by race/ethnicity. However: (i) on all sentencing margins, statistically significant Black-White differentials remain; (ii) for Hispanic-White differentials, there is no difference in the likelihood to be downward departed, but the conditional sentencing gap opens up of 3.7 months, that is not statistically different from the Black-White sentencing gap [$p = .607$].

To assess whether these conditional differentials can be attributed to unobserved heterogeneity, we follow Altonji *et al.* [2005] and Oster [2017] to estimate *bounds* on the treatment effect of race/ethnicity allowing for selection on unobservables (SoU). Key to this is making an assumption on how the unobserved and observed covariates driving outcomes relate to each other. Altonji *et al.* [2005] and Oster [2017] assume they relate through a proportional selection relationship where the coefficient of proportionality is denoted τ. It can then be shown that the true causal impact for ethnic group e, δ_e^*, depends on δ (and other factors): $\delta_e^* = \delta_e(\tau, .)$. Bounds on δ_e are such that, at one extreme, if $\tau = 0$ the unobserved covariates do not bias the conditional specification and $\delta_e^* = \delta_e$. At the other extreme, equal selection ($\tau = 1$) is an appropriate upper bound on τ: intuitively, the set of unobservables cannot be *more* important than the available covariates in explaining the treatment effect of race/ethnicity on outcomes. This is plausible in our context given we observe a rich set of defendant and legal characteristics including the recommended guideline cell, that lead to a relatively high R^2 as reported in Table 1. The bounds reported in Table 1 are $\delta_e(0) = \delta_e$ and $\delta_e(1)$, and we also report the τ required for $\delta_e(\tau) = 0$.

Columns 1 to 3 show that allowing for SoU: (i) the Black-White bounds include zero along all margins. For example, for there to be no Black-White differential on the sentence length margin, $\tau = .341$ is required. In contrast, for Hispanic-White sentencing differentials there remains evidence of a gap in sentence length: $\delta_e \in [3.67, 5.17]$ and $\delta_e(\tau) = 0$ requires $|\tau| > 2$, so unobservables would need to be *more* than twice as important in explaining the Hispanic-White differential than the observables conditioned on, including the recommended guideline cell, and these unobservables should be opposite correlation to ethnicity on the any sentence and sentence length margins.

To be clear, this does not rule out there being discrimination against Black male defendants in the Federal CJS. Rather the estimated bounds highlight that conditional Black-White differences could go to zero if unobservable characteristics of Black defendants driving sentencing outcomes are correlated to their observed covariates to a plausible degree ($\tau \leq 1$). This is not the case for Hispanic-White sentencing gaps, for which the evidence suggests can only by ruled out by the omission of covariates under even more implausible conditions (the sign of τ varies across sentencing margins and $|\tau| > 1$).[6]

Columns 4 to 6 repeat the analysis for female defendants. We have a large sample, and this is useful given the majority of research on the CJS focuses on men. For women defendants we see a pattern of results broadly similar to those for men. In particular, there are large unconditional sentencing disparities across race/ethnicity for women, although the magnitudes tend to be smaller than for men. As for men, these differentials are robust to conditioning on observables. When doing so, the Black-White (Hispanic-White) sentencing gap is 1.4 (1.8) months, that are both statistically significant, and around half the size of male defendant gaps on the same margin. Accounting for SoU we see that Black-White and Hispanic-White sentencing margins remain robust on the any sentence and sentence length margins. As for men, there is little evidence of racial/ethnic sentencing differentials along the margin of downward departures.

4 Conclusions

It is hard to design policy responses to address racial/ethnic sentencing gaps without understanding the root cause of such gaps. The central challenge lies in understanding whether much documented differential outcomes by race/ethnicity are driven by unobserved heterogeneity across defendants, or whether they reflect true discrimination. Our analysis highlights that using rich data on Federal cases, accounting for SoU suggests that for: (i) for male defendants, Hispanic-White sentencing

[6]The fact that the SoU bounds on downward departures include zero for Black-White differences is important in light of earlier research, including Mustard [2001], documenting the majority of the Black-White sentencing differential is attributable to differences in downward departure.

gaps are more robust than Black-White gaps, although for neither group do we find robust evidence of gaps in downward departures; (ii) for female defendants, Black-White and Hispanic-White sentencing length gaps are robust to SoU, although again for neither group do we find robust evidence of gaps in downward departures. The future challenge lies in setting up research designs, as in Anwar [2006], Abrams *et al.* [2012], Anwar *et al.* [2012] and McConnell and Rasul [2018], that go beyond this and measure causal estimates of race/ethnicity on sentencing outcomes.

References

[1] ABRAMS.D, M.BERTRAND AND S.MULLAINATHAN (2012) "Do Judges Vary in their Treatment of Race?," *Journal of Legal Studies* 41: 347-83.

[2] ALTONJI.J.G, T.E.ELDER AND C.R.TABER (2005) "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* 113: 151-84.

[3] ANWAR.S AND H.FANG (2006) "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," *American Economic Review* 96: 127-51.

[4] ANWAR.S, P.BAYER AND R.HJALMARSSON (2012) "The Impact of Jury Race in Criminal Trials," *Quarterly Journal of Economics* 127: 1017-55.

[5] MCCONNELL.B AND I.RASUL (2018) Contagious Animosity in the Federal Criminal Justice System, mimeo UCL.

[6] MUSTARD.D.B (2001) "Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the US Federal Courts," *Journal of Law and Economics* 44: 285-314.

[7] NEAL.D AND A.RICK (2015) "The Prison Boom and Sentencing Policy," forthcoming, *Journal of Legal Studies.*

[8] OSTER.E (2017) "Unobservable Selection and Coefficient Stability: Theory and Validation," forthcoming *Journal of Business Economics and Statistics.*

[9] SORENSEN.T, S.SARNIKAR AND R.L.OAXACA (2012) "Race and Gender Differences Under Federal Sentencing Guidelines," *American Economic Review, Papers & Proc.* 102: 256-60.

[10] USSC (1999-2003) Monitoring of Federal Criminal Sentences, 1999-2003 [Computer file], ICPSR version. Wash. DC: USSC [producer], 1999-2006. Ann Arbor, MI: ICPSR [distrib.].

Table 1: Ethnic Sentencing Differentials, by Gender

Sample: Federal Cases up for Sentencing between 10/1/1998 and 9/30/2003
Standard errors in parentheses clustered by ethnicity-district
Selection on unobservables (SoU) bounds in brackets

		Men			Women		
		(1) Downward Departure	(2) Any Sentence	(3) Sentence Length	(4) Downward Departure	(5) Any Sentence	(6) Sentence Length
Black	**Unconditional**	-.047**	.120***	40.3***	-.045*	-.006	3.33*
		(.019)	(.012)	(3.44)	(.027)	(.024)	(1.75)
	Conditional	-.010*	.028***	3.97***	-.013**	.008*	1.35***
		(.006)	(.004)	(.468)	(.005)	(.005)	(.310)
	[Bounds: $\delta_{B(0)}$, $\delta_{B(1)}$]	[-.010, .003]	[-.003, .028]	[-7.91, 3.97]	[-.013, -.002]	[.008, .012]	[.753, 1.35]
	τ required for coefficient of 0	.783	.903	.341	1.138	-1.823	2.259
Hispanic	**Unconditional**	.122**	.165***	-.833	.097*	.159***	3.45*
		(.060)	(.011)	(4.05)	(.058)	(.022)	(2.07)
	Conditional	.002	.059***	3.69***	.012	.060***	1.81***
		(.015)	(.006)	(.487)	(.008)	(.009)	(.378)
	[Bounds: $\delta_{H(0)}$, $\delta_{H(1)}$]	[-.041, .002]	[.023, .059]	[3.69, 5.17]	[-.017, .012]	[.029, .06]	[1.31, 1.81]
	τ required for coefficient of 0	.050	1.54	-2.45	.419	1.85	3.63
Sentencing Outcome for Whites		.122	.784	42.6	.138	.608	19.6
Unconditional p-value: [Black = Hispanic]		.003	.000	.000	.011	.000	.956
Conditional p-value: [Black = Hispanic]		.267	.000	.607	.001	.000	.243
R^{max}=min(1, 1.3 x unadjusted R-squared)		.306	.565	.977	.283	.610	.967
Conditional Adjusted R-squared		.234	.434	.751	.209	.464	.741
Observations		235,484	235,484	235,484	38,690	38,690	38,690

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in all Columns. Standard errors are reported in parentheses, where these are clustered by ethnicity-district. The full sample of Federal cases is used (those that come up for sentencing from 10/1/1998 to 09/30/2003), totaling 235, 484 for male defendants and 38, 690 for female defendants. The dependent variable in Columns 1 and 4 is a dummy for whether the case receives a downwards departure. The dependent variable in Columns 2 and 5 is a dummy for receiving a prison sentence. The dependent variable in Columns 3 and 6 is the sentence length (in months) including zero. For unconditional estimates, we only condition on defendant ethnicity (White, Black, Hispanic). For conditional estimates, the following additional controls are included: fiscal year dummies, on offender characteristics, we control for dummies for the highest education level, marital status, a dummy for whether age is missing, age and age squared interacted with this non-missing age dummy, a dummy for whether the number of dependents is missing, and the number of dependents interacted with a non-missing dependents dummy; on legal controls, we control for a dummy whether information on the defense counsel is missing, and a non-missing dummy interacted with the type of defense counsel (privately retained, court appointed, federal public defender, self-represented, rights waived, other arrangements); the primary offense type, the guideline cell, and Federal district dummies. The p-value at the foot of each Column is on the null that the coefficients on the Black and Hispanic dummy are equal against a two sided alternative. In parentheses we report bounds on the OLS estimate accounting for selection on unobservables using the Oster [2017] method: the bounds are set assuming the coefficient of proportionality is zero or one. Below the bounds we report the coefficient of proportionality that is required for the implied point estimate to be zero.

Table A1: Sentencing Guideline Cells

(months imprisonment)

		Criminal History Category (Criminal History Points)					
		I	II	III	IV	V	VI
		(0 or 1)	(2 or 3)	(4, 5, 6)	(7, 8, 9)	(10, 11, 12)	(13 or more)
	1	0-6	0-6	0-6	0-6	0-6	0-6
	2	0-6	0-6	0-6	0-6	0-6	1-7
	3	0-6	0-6	0-6	0-6	2-8	3-9
	4	0-6	0-6	0-6	2-8	4-10	6-12
Zone A	5	0-6	0-6	1-7	4-10	6-12	9-15
	6	0-6	1-7	2-8	6-12	9-15	12-18
	7	0-6	2-8	4-10	8-14	12-18	15-21
	8	0-6	4-10	6-12	10-16	15-21	18-24
Zone B	9	4-10	6-12	8-14	12-18	18-24	21-27
	10	6-12	8-14	10-16	15-21	21-27	24-30
Zone C	11	8-14	10-16	12-18	18-24	24-30	27-33
	12	10-16	12-18	15-21	21-27	27-33	30-37
	13	12-18	15-21	18-24	24-30	30-37	33-41
	14	15-21	18-24	21-27	27-33	33-41	37-46
	15	18-24	21-27	24-30	30-37	37-46	41-51
	16	21-27	24-30	27-33	33-41	41-51	46-57
	17	24-30	27-33	30-37	37-46	46-57	51-63
	18	27-33	30-37	33-41	41-51	51-63	57-71
	19	30-37	33-41	37-46	46-57	57-71	63-78
	20	33-41	37-46	41-51	51-63	63-78	70-87
	21	37-46	41-51	46-57	57-71	70-87	77-96
	22	41-51	46-57	51-63	63-78	77-96	84-105
	23	46-57	51-63	57-71	70-87	84-105	92-115
	24	51-63	57-71	63-78	77-96	92-115	100-125
	25	57-71	63-78	70-87	84-105	100-125	110-137
	26	63-78	70-87	78-97	92-115	110-137	120-150
	27	70-87	78-97	87-108	100-125	120-150	130-162
Zone D	28	78-97	87-108	97-121	110-137	130-162	140-175
	29	87-108	97-121	108-135	121-151	140-175	151-188
	30	97-121	108-135	121-151	135-168	151-188	168-210
	31	108-135	121-151	135-168	151-188	168-210	188-235
	32	121-151	135-168	151-188	168-210	188-235	210-262
	33	135-168	151-188	168-210	188-235	210-262	235-293
	34	151-188	168-210	188-235	210-262	235-293	262-327
	35	168-210	188-235	210-262	235-293	262-327	292-365
	36	188-235	210-262	235-293	262-327	292-365	324-405
	37	210-262	235-293	262-327	292-365	324-405	360-life
	38	235-293	262-327	292-365	324-405	360-life	360-life
	39	262-327	292-365	324-405	360-life	360-life	360-life
	40	292-365	324-405	360-life	360-life	360-life	360-life
	41	324-405	360-life	360-life	360-life	360-life	360-life
	42	360-life	360-life	360-life	360-life	360-life	360-life
	43	life	life	life	life	life	life

Offense Level

Source: Chapter 5, 2001 Federal Sentencing Guidelines Manual [http://www.ussc.gov/sites/default/files/pdf/guidelines-manual/2001/manual/CHAP5.pdf]

Table A2: Descriptives, by Ethnicity and Gender

Means, standard deviations in parentheses.

	Men			Women		
	White	**Black**	**Hispanic**	**White**	**Black**	**Hispanic**
Sample Size	73,786	60,653	101,045	15,686	12,076	10,928
Number Dependents	1.15	1.68	1.85	1.11	1.54	1.76
	(1.43)	(1.84)	(1.80)	(1.30)	(1.54)	(1.64)
Age	38.6	31.9	32.1	36.6	33.2	33.0
	(12.2)	(9.29)	(9.22)	(11.3)	(9.94)	(10.1)
Marital Status:						
Single	.337	.536	.328	.275	.504	.335
Married	.353	.205	.344	.341	.189	.245
Other	.289	.237	.244	.363	.275	.355
Education Level:						
Less than High School	.260	.404	.635	.237	.295	.533
High School Graduate	.377	.365	.165	.409	.338	.239
Some College	.226	.184	.072	.267	.296	.130
College Graduate	.125	.038	.019	.073	.056	.031
Defense Counsel:						
Privately Retained	.169	.080	.072	.128	.071	.103
Court Appointed	.173	.176	.298	.202	.197	.266
Federal Public Defender	.122	.141	.262	.124	.150	.199
Other	.007	.006	.002	.010	.013	.004
Criminal History Score	2.19	3.05	2.40	1.57	1.71	1.41
	(1.64)	(1.84)	(1.67)	(1.16)	(1.30)	(1.00)
Offense Severity	17.9	22.2	18.4	14.6	14.5	16.4
	(8.45)	(9.43)	(8.05)	(7.91)	(8.39)	(7.67)

Notes: The full sample refers to all Federal cases that come up for sentencing from 10/1/1998 to 09/30/2003. For each gender we show the descriptive statistic for each race/ethnicity. Standard errors are shown in parentheses for continuous variables. For marital status, the "Other" category incorporates cohabiting, divorced, widowed and separated. For defense counsel the "Other" category incorporates defendant represented self, waived rights to counsel and other arrangements for counsel.