Hispanic-White Sentencing Differentials in the Federal Criminal Justice System[*]

Brendon McConnell   Imran Rasul

June 2016

Abstract

In the Federal criminal justice system, large differences in sentencing outcomes exist between Hispanic and White defendants. A candidate explanation is ingroup bias causing 'outsiders' (Hispanics) to be treated differently to 'insiders' (Whites). To probe this explanation we exploit 9-11 as an exogenously timed cue heightening the salience of insider-outsider differences in American society. Based on linked administrative data that covers criminal cases from time of arrest through to sentencing, we use a DiD research design based on defendants *all* of whom were arrested pre 9-11, but some were sufficiently far advanced along the timeline so as to come up for sentencing pre 9-11, while others had only just entered the timeline prior to 9-11, and so were sentenced post 9-11. We document that among those sentenced post 9-11, Hispanic-White judicial sentencing differentials are further exacerbated relative to these sentenced pre 9-11, while Black-White sentencing differentials are unaffected. Our linked administrative data and research design also allows us to document the differential treatment of Hispanic defendants by prosecutors in pre-sentencing stages of the CJS, such as with regards to the initial offense charges they set. Finally, we collate bibliographical information on judges and document that in districts with a higher proportion of Hispanic judges, the Hispanic-White sentencing differential is significantly reduced, consistent with judges' ingroup biases driving their sentencing decisions. Our results provide insights into the magnitude, channels and potential origins of Hispanic-White sentencing differentials in the Federal criminal justice system. *JEL Classification: J15, K14.*

1 Introduction

Ethnic minority men are far more likely to come into contact with the Federal criminal justice system (CJS) than White men. Decades of research have shown sentencing outcomes also vary by ethnicity. While much of this has focused on Black-White differences, Hispanics are: (i) the modal defendant in the Federal CJS and are four times as likely to go to prison during their lives as Whites [Starr and Rehavi 2013]; (ii) the group whose incarceration rate is growing fastest; (iii) a minority group that have been somewhat understudied relative to other minorities in the economics of discrimination literature more generally [Charles and Guryan 2011], despite Hispanics being ever more prominent in the political, legal and cultural life of America.

The challenge lies in understanding whether sentencing differentials are driven by unobserved heterogeneity across defendants, correlated to their ethnicity, or whether they reflect discrimination. The question is important given that equality before the law is a cornerstone of any judicial system, and because it is difficult to know what could be done to reduce sentencing disparities if their underlying causes remain unknown. Our key contribution is to combine linked administrative data with a research design to provide insights into the magnitude, channels and potential origins of Hispanic-White sentencing differentials in the Federal criminal justice system. We build an evidence base that is consistent with these differentials being driven by the ingroup biases of judges and prosecutors, whereby 'outsiders' (Hispanics) are treated differently to 'insiders' (Whites).

We use the *Monitoring of Federal Criminal Sentences* (MFCS) data set: this comprises information from four linked administrative data sources covering a defendant's arrest, and all subsequent stages of their processing through the Federal CJS shown in Figure 1. A prominent set of papers have used State or lower court data to estimate the causal impact of sentence length on criminal and economic outcomes. These have exploited the random assignment of judges to cases, so that variation in the harshness of judges leads to exogenous variation in the sentence length received by defendants [Kling 2006, Abrams *et al.* 2012, Aizer and Doyle 2015, Mueller-Smith 2016]. In Federal court data, even though judges are randomly assigned to cases, because criminal cases considered are more serious and often of national importance, judge identifiers are typically unavailable. The key advantages of Federal criminal court data however relate to tackling long-standing challenges on empirical work on the CJS [Klepper *et al.* 1983]: (i) it is nationally representative, covering cases from all 90 mainland US Districts, defendants of all ages, and all types of criminal offense; (ii) the linked administrative data allows pre-sentencing differential treatment arising from the behavior of prosecutors or legal counsel to be studied alongside the behavior of judges; (iii) large samples are available allowing for both Black-White and Hispanic-White differentials to be studied: our working sample covers 230, 000 Federal criminal cases occurring between 1998 and 2003.

The research design we utilize allows us to make progress on understanding the magnitude, channels and potential origins of Hispanic-White sentencing differentials in the US Federal CJS. A candidate explanation for such sentencing differentials is that ingroup biases lead 'outsiders' to US society (Hispanics) being treated differently to 'insiders' (Whites). There is a vast literature examining the biological and evolutionary roots of ingroup bias. Ingroup biases have been documented in lab and field settings [Shayo 2009, Bertrand and Duflo 2016]. There is also evidence for ingroup bias in judicial contexts including US State courts [Bushway and Piehl 2001, Shayo and Zussman 2011, Abrams *et al.* 2012, Anwar *et al.* 2012, Rehavi and Starr 2014]. Our analysis sheds new light on whether such biases drive the Hispanic-White sentencing differentials in the high stakes environment of the Federal criminal justice system, where decisions are made by professional and experienced judges, prosecutors and legal counsel, and the universe of criminal offenses and district courts can be studied.

We consider 9-11 as an exogenously timed event that heightened the salience of insider-outsider differences in US society and so might have cued pre-existing ingroup biases [Woods 2011]. To isolate the impact the event had on sentencing outcomes in the Federal CJS, we compare sentencing outcomes between: (i) defendants who committed their last offense before 9-11 and were sentenced *before* 9-11 (the control group); (ii) defendants who *also* committed their last offense before 9-11, but were sentenced *after* 9-11 (the treated group). We construct a second difference in outcomes across ethnicities to estimate a difference-in-difference (DiD) impact of 9-11 on sentencing. We base our sample on a ±180 day sentencing window around 9-11 2001, where *all* defendants have committed their offense prior to 9-11, and hence entered Stage 1 of the Federal CJS timeline in Figure 1, but some were sufficiently far advanced along the timeline so as to come up for sentencing pre 9-11, while others had only just entered the timeline prior to 9-11, and so ended up being sentenced post 9-11. The specific advantage of exploiting 9-11 as a natural experiment is that it likely cues the kinds of ingroup bias that might drive Hispanic-White sentencing differentials.

The period we study is when sentencing guidelines are in place in the Federal CJS, where these guidelines provide for *determinate* sentencing. Table A1 shows the full set of guideline cells, mapping combinations of the severity of the offense and the defendant's criminal history into a specific sentencing range. The guidelines do however allow judge's discretion to *downwards depart* from the recommended guideline cell, and so move in a North-Westerly direction in Table A1. Our analysis of sentencing outcomes considers ethnic differentials in the guideline cell into which guilty defendants are placed, total sentence length, and sentence movements within guideline cells.

We first confirm that relative to Whites, Hispanics sentenced pre 9-11 are more likely to go to prison, and receive longer sentences on average and at the median. For Hispanics sentenced post

9-11 when the salience of insider-outside divisions is heightened, sentencing differentials become even further exacerbated: Hispanics become 13.5% less likely to receive a downward departure than Whites, and their median sentence lengths significantly rise by .632 months, suggesting that even within a guideline cell, Hispanics are sentenced more harshly despite these defendants having committed their crime pre 9-11.

Black-White sentencing differentials around 9-11 are unaffected along all sentencing margins.

To underpin a causal interpretation of these results, we provide evidence for two identifying assumptions. We first show the time a defendant spends in the CJS between when their last offense is committed and when they come up for sentencing is not impacted by 9-11. Second, using data from other years to construct placebo 9-11 impacts, we show there are no ethnicity-time effects in ethnic sentencing differentials that occur naturally around 9-11 each year.

As has long been recognized [Klepper *et al.* 1983] a range of legal actors beyond judges are involved in the Federal CJS, and their behaviors: (i) can lead to differential treatment by ethnicity pre-sentencing; (ii) such differential treatment might not be detected in sentencing differentials. These concerns are heightened when sentencing guidelines are in place as these restrict the discretion of judges and might increase the power of prosecutors [Starr and Rehavi 2013]. We address the issue by combining the linked administrative data with our research design to consider decisions made at earlier stages of the case timeline on Figure 1, where we move our 9-11 window to when these other decisions are being made. We examine: (i) prosecutor decisions at the sentencing stage in asking for substantial assistance departures; (ii) prosecutor-legal counsel interactions in drafting pre-sentence reports that provide judges with a recommended guideline cell; (iii) prosecutor decisions over which initial offense charge to file. Echoing the findings of Rehavi and Starr [2014], we find that post 9-11, Hispanic defendants are significantly more likely to receive an initial offense that carries a statutory minimum, the statutory minimum sentence they receive on average is 11 months longer, and this translates into Hispanics eventually receiving sentences that are 9 months longer than Whites. Given most Federal prosecutors are not Hispanic, we interpret this pattern of prosecutor's initial offense charge decisions to be consistent with them displaying ingroup biases.

In summary, our data and research design allows us to pinpoint two cohorts of Hispanic defendant whose treatment in the Federal CJS appears subject to ingroup biases: (i) the cohort that come up for sentencing just post 9-11, who are significantly less likely than Whites to receive a downwards departures from judges; (ii) the cohort whose initial offense charges are set by Federal prosecutors post 9-11, who are charged with offenses with significantly longer statutory minimum sentences. For both cohorts, a Juhn-Murphy-Pierce decomposition of the DiD in sentencing outcomes shows these differences are largely driven by changes in unobserved drivers of

sentencing outcomes; only negligible amounts of each cohort's unconditional DiD in outcome can be attributed to either the DiD in their observables relative to Whites, or the sentencing prices of such observables (such as immigration offenses being treated more harshly post 9-11).

Finally, we return to consider ingroup biases of *judges* by analyzing how judge characteristics correlate to our estimated Hispanic-White sentencing differential. To make progress on the issue we have hand-coded characteristics of Federal judge's by district court, sourced from the *Biographical Directory of Federal Judges*. We document that in districts where there is a higher proportion of Hispanic judges, the Hispanic-White sentencing differential for downward departures is significantly reduced. This is in line with judges displaying ingroup bias [Schanzenbach 2005, Abrams *et al.* 2012]. We find post 9-11 Hispanic-White sentencing differences for downward departures are generally more pronounced in those districts where there is a smaller share of Hispanics in the population. Such a spatial pattern of sentencing differentials is in line with the contact hypothesis, that states that interpersonal contact is an effective ways to reduce prejudice between majority and minority group members [Allport 1954].

To examine the external validity of these findings, we repeat the exercise but based on Hispanic-White differentials for downward departures in the full sample of 230,000 Federal criminal cases from 1998 to 2003. Strikingly, we continue to see evidence of ingroup bias that is of a similar magnitude as what is found based on the natural experiment sample estimates.

Our contribution is to combine linked administrative data with a novel research design to study Hispanic-White sentencing differentials. We show the presence of ingroup biases determine sentencing outcomes in the high stakes and professional environment of the Federal CJS, and advance the literature by pinpointing the separate roles that judges and prosecutors have in driving the differential treatment of Hispanic defendants in the Federal CJS [Shayo and Zussman 2011, Abrams *et al.* 2012, Anwar *et al.* 2012, Rehavi and Starr 2014].

The paper is organized as follows. Section 2 describes the Federal CJS, sentencing guidelines, and administrative data. Section 3 presents our DiD research design using 9-11 to shed light on whether ethnic sentencing differentials might originate from ingroup biases. Section 4 presents our findings on sentencing outcomes determined by judges. Section 5 focuses on pre-sentencing stages to narrow down the role that other actors play in driving differential treatment pre-sentencing. Section 6 examines the cross-district evidence for ingroup biases of judges. Section 7 concludes. The Appendix contains further data details and robustness checks.

4

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. If both Federal and State courts have jurisdiction over a criminal act, Federal and State prosecutors make case-by-case decisions as to which court the defendant will be tried in. The sorting of cases into systems is therefore an executive branch decision: judges and defense counsel have no formal role. The DiD research design we use to estimate Hispanic-White sentencing differentials eliminates time invariant cross sectional differences between defendants of different ethnicity among those tried in the Federal system.

As criminal cases heard in Federal courts tend to be more serious than those in State courts, the types of offense considered differ from those in State courts. For example, in 2000 the three most frequent criminal offenses filed in Federal courts were for drug trafficking (40%), immigration (22%), and fraud (9%), while at the State level the most frequent criminal cases related to drug sales (19%), other drug offenses (18%), and assault (10%). Sentencing severity is harsher in Federal court: 88% (75%) of those convicted in Federal (State) court receive a custodial sentence, with the mean sentence being 67 (48) months in Federal (State) court.[1]

The legal actors determining sentencing outcomes in Federal criminal cases are judges, prosecutors, the defendant's legal counsel, and juries. Judges in Federal courts are nominated by the President, confirmed by Congress, and appointed for life (in contrast, State court judges can be elected, appointed or a combination). There are just over 7 Federal judges per district, so that there are around 700 in total: they are among the most senior judges in the country, and *a priori*, might be considered among those least susceptible to biased judgments. The prosecution of Federal criminal cases in each of the 94 US District Courts is the responsibility of the US Attorney for that District, who is also a Presidential appointee reporting directly to the Attorney General.

Legal counsel in Federal courts differs from State courts: in 47% of Federal criminal cases, legal counsel is court appointed. Federal public defenders operate in 32% of cases, and 21% of defendants retain private counsel. This differs from State court cases where 68% of defendants have a public defender. Finally, jury trials in Federal courts occur only if a defendant pleads not guilty. In the Federal CJS 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, and such pleas can be Pareto improving for risk averse defendants and prosecutors. By pleading guilty, defendants give up the right to appeal except in capital cases (that represent less than .1% of cases) [Alesina and La Ferrara 2015].

---

[1] The difference in severity across courts is not driven by the composition of offenses: within offense type there is considerably harsher sentencing in Federal courts, reflecting the greater seriousness of such crimes.

2.1 Timeline

Figure 1 shows the timeline of Federal criminal cases. Table A2 further details each stage. The first stage a defendant faces after having been arrested and formally charged with a Federal offense (Stage 0) is their initial court appearance where their defense counsel is assigned (Stage 1). Bail is then determined (Stage 2), initial charges are filed by prosecutor's during arraignment (Stage 3), leading to the defendant's initial district court appearance (Stage 4), where they find out which judge they have been assigned to. Pre-trial motions take place at Stage 5, to determine what evidence can be used in trial. The defendant can then offer a plea (Stage 6), where 96% plead guilty. The trial represents Stage 7, and sentencing occurs at Stage 8. In rare cases where a defendant pleads not guilty or for capital cases, they retain the right to appeal (Stage 9).

We first focus on sentencing (Stage 8), given this is where judges exercise their discretion over defendant outcomes, and as 96% of defendants are already convicted, only their punishment remains to be determined. The ethnic sentencing differentials we measure are conditional on defendant's reaching sentencing Stage 8. This includes conditioning on the guideline cell recommended to the judge in the pre-sentence report drawn up by the defense counsel and prosecutor between trial and sentencing. Multiple legal actors are involved at earlier stages, and: (i) their behaviors can lead to differential treatment of defendants pre-sentencing; (ii) the presence of biases earlier in the timeline might not be detected in judicial sentencing differentials. This might especially be so when sentencing guidelines are in place as these restrict the discretion of judges and potentially increase the power of prosecutors [Starr and Rehavi 2013]. In Section 5 we exploit the linked administrative data to consider earlier stages to pin point the influence of other legal actors in driving ethnic sentencing differentials, including the initial offense charges of prosecutors that have been shown to play an important role in driving Black-White sentencing gaps [Rehavi and Starr 2014]. By analyzing the decision making of multiple actors, a novel aspect of our analysis is that it allows us to measure whether the sentencing behavior of Federal judges reinforces or offsets the behavior of other legal actors with regards to Hispanic-White sentencing differentials.

2.2 Linked Administrative Data

We use the *Monitoring of Federal Criminal Sentences* (MFCS) data set for our analysis. This comprises information gathered from four linked administrative data sources covering the arrest/offense stage before an individual enters the Federal CJS (Stage 0), and all subsequent stages shown in Figure 1. We focus on male defendants so our 'full sample' covers 250,000 Federal criminal cases that come up for sentencing from October 1998 to September 2003 across nearly all US districts [USSC MCFS 1999-2003]. The Appendix provides further details on the data.

To estimate Hispanic-White and Black-White sentencing differentials, we use two variables available at the sentencing Stage 8 in the *MFCS* data, to code each defendant's ethnicity. In one variable, defendants are classed as either Hispanic (41%) or non-Hispanic (59%). A separate race code then separately identifies defendants as white-race (71%), black-race (29%), other-race ($< .1\%$). Combining these variables, Whites are coded as white-race and non-Hispanic; Blacks are coded as black-race and non-Hispanic; Hispanics are coded as white- or black-race and Hispanic. This coding implies 31% of defendants are ethnically White, 26% are Black and 43% are Hispanic. The large sample of criminal cases in the *MCFS* data allows us to study both Hispanic-White and Black-White sentencing differentials.[2]

The *MFCS* data contains a rich set of 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 other pre-sentence variables (such as whether the defendant is in custody), and offense details are recorded that we use to classify the offense into various types (drug, immigration etc.). Most importantly, the data records the guideline cell recommended to the judge in the pre-sentence report. This effectively proxies all case-specific factors the prosecution and legal counsel deem judges should factor into their sentencing decision. Finally the data record the Federal court district of sentencing. Table A3a shows the sample descriptives for the *MCFS* full sample of cases, as well as the working sample we use for our analysis based on the 94% of cases in which there is no missing data on the core covariates.

2.3 Linkage Rates

A concern when studying sentencing outcomes is that there can be selection of defendants into this stage of the CJS [Klepper *et al.* 1983]: as the result of actions of various legal actors through the case timeline, the set of cases that reach sentencing might not be representative of the original population of arrested and charged defendants. As the *MCFS* data comprises linked administrative sets covering arrest/offense Stage 0 through to sentencing Stage 8, we can estimate linkage rates for criminal cases across stages. We first consider cases observed at sentencing Stage 8, and estimate linkage rates to the *earlier* administrative records, as shown in Panel A in the lower part of Figure 1 (right-to-left linkage rates). To prevent linkage rates being spuriously lowered due to case truncation, we consider cases up for sentencing in the final year of our *MFCS* data. We see that: (i) 90% of cases are also observed in the preceding administrative data (covering

---

[2]The other-race classifications include American Indian/Alaskan Native, Asian/Pacific Islander, multi-racial and other. The *MFCS* data thus does not contain an identifier for Arabs nor Muslims, and so those groups are not the focus of our study (even if such identifiers existed, the numbers of such defendants would be miniscule, corresponding to less than .1% of criminal cases). Using our coding, 92% of Hispanics are white-race.

Stages 4-7); (ii) 85% of cases observed at sentencing can be further linked back to the two earlier administrative data sets (covering Stages 1-7); (iii) 75% of cases observed at sentencing can be linked back to arrest/offense stage. Linkage rates are quite similar across ethnicities: 72% of records for White defendants up for sentencing can be linked all the way back to the arrest/offense stage; the corresponding rates for Black (Hispanic) defendants are 70% (81%). For drug offenses linkage rates back to the arrest/offense stage are 74-78% across ethnicities, and for immigration offenses they are 71-85%. The fact that linkage rates are less than 100% implies either: (i) truncation of cases because some cases started before 1998 (our first year of data); (ii) linkage errors arising from the fact the *MCFS* data originates from multiple agencies.

We next construct linkage rates from the arrest/offense stage through to sentencing, as shown in Panel B in the lower part of Figure 1 (left-to-right linkage rates). The drawback is that only race is coded in the arrest/offense Stage 0 so when deriving these linkage rates we can only do so for white-race and black-race defendants (92% of those coded as Hispanic at sentencing are white-race). To again minimize linkage rates being spuriously lowered due to truncation, we consider cases where arrest/offense dates occur in the first year of our *MCFS* data. The underlying administrative set from which the arrest/offence data are collected is from the US Marshals Service data, and this includes all persons arrested by Federal law enforcement agencies, persons arrested by local officials and then transferred to Federal custody, and persons who avoid arrest by self-surrendering. Around 38% of such individuals actually enter the Federal CJS at Stage 1, and this rate is similar for white- and black-race individuals (38-39%). These rates might reflect that in the majority of cases, either prosecutors do not pursue any case at all or that individuals are assigned to be tried in State courts. We see higher linkage rates for drug offenses, that do not vary much by race (54-55%), but for immigration offenses, black-race individuals are more likely to enter the Federal CJS (45% versus 34%). Most importantly though, once an individual enters the Federal CJS at Stage 1, there remains a high linkage rate to the *subsequent* administrative data sets: (i) 84% of defendants in Stage 1 can be traced though to Stage 8; (ii) linkage rates are similar across races (84-86%), and across races for drug offenses (86-88%) and immigration offenses (76-82%).

To reiterate, the difference-in-difference research design we utilize to estimate ethnic sentencing differentials eliminates time invariant cross sectional differences between defendants of different ethnicity among those assigned to be tried in the Federal system.

2.4   Federal Sentencing Guidelines

Federal sentencing guidelines were introduced in the Sentencing Reform Act of 1984 by the US Sentencing Commission (USSC). The explicit goal of the reform was to alleviate sentencing dis-

parities that research had indicated were prevalent in the Federal CJS. This was to be achieved by the guidelines providing for *determinate* sentencing, whereby: (i) the discretion judges had over penalties imposed at the sentencing stage became more limited; (ii) parole boards were abolished so that determined sentences matched the actual period of incarceration far more closely.[3]

The USSC sentencing guidelines are based on: (i) the severity of the offense; (ii) the defendant's criminal history. To run through a stylized example, an individual who commits a robbery is allocated a base level of 20 points. If a gun is involved an additional 5 points are awarded (if the individual had been a minimal participant in the robbery, 4 points would have been deducted). If the individual was found to be in obstruction of justice, an additional 2 points are awarded. Hence in this case the final score of the defendant on offense severity would be 23 points. There are six criminal history categories, each associated with a range of criminal history points. 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.

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. These include those in Zone A on Table A1, where the guidelines include zero sentence length, and cells in Zone D where the guidelines impose a life sentence. Accounting for the empirical distribution of offense severity and criminal histories, the expected width of a guideline cell is 15 months, and the sentencing range within a guideline cell therefore corresponds to around 25% of the minimum sentence [Schanzenbach 2005]. Our analysis of sentencing outcomes considers ethnic differentials in the guideline cell into which guilty defendants are placed, total sentence length, and sentence movements within guideline cells.

Between trial/conviction and sentencing (Stages 7 and 8), the pre-sentence report is drafted by prosecutors and legal counsel, and this specifies a recommended guideline cell. However, the sentencing guidelines still provide *judges* discretion over which guideline cell to ultimately place a defendant in. They allow a judge to *downwards depart* from the recommended guideline cell, and so move in a North-Westerly direction in the guideline cell Table A1. A judge can do so

---

[3]This is in contrast to the prior system of indeterminate sentencing, in which a sentence with a maximum (and perhaps a minimum) was pronounced by a judge, but the actual time served in prison was determined by a parole commission after the sentence began. As part of the same reforms, such parole on Federal cases was abolished. The notion that the majority of a Federal court sentence should be served is also something that has become strengthened by other Federal laws, such as truth-in-sentence (TIS) laws, that further eliminate or restrict parole and/or remissions. In 1994, a Federal TIS law stated that to qualify for TIS Federal funding, offenders must serve at least 85% of the sentence for qualifying crimes before becoming eligible for parole. As of 2008, 36 states qualified for this additional funding.

if they find mitigating circumstances of a kind not adequately taken into consideration by the USSC in formulating the sentencing guidelines. These circumstances 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.[4]

In our full sample of 230,000 Federal criminal cases from October 1998 to September 2003, judges grant downwards departure in 17% of cases. Downward departures result in a sentence below the original guideline range but they still lead to a custodial sentence in almost 90% of cases. Upwards departures are permitted but occur in less than 1% of cases.

Judge-initiated downwards departures are a key sentencing outcome to consider because: (i) such decisions are cleanly attributable to judges; (ii) they are typically associated with reductions in sentence length; (iii) they are likely correlated to the prison conditions under which incarceration is served, and this in turn might impact recidivism and other future behaviors through the accumulation of criminal capital [Bayer *et al.* 2009]. The null hypothesis for our analysis is based on the USSC sentencing guidelines themselves that state that *"race, sex, national creed, religion and socioeconomic status"*, are factors that *"are not relevant in the determination of a sentence"* *[§5H1.10 of the sentencing guidelines]*.[5]

3   Ethnic Sentencing Differentials

3.1   Preliminary Evidence

We consider four margins of judicial sentencing: (i) if a downward departure is granted; (ii) if any prison sentence is given; (ii) the sentence length (in months) including zero for non-custodial sentences; (iv) the median sentence length. In the full sample of 230,000 Federal criminal cases, large sentencing differentials exist between ethnicities. Panel A of Table 1 shows that relative to Whites: (i) Blacks are 4.7pp less likely to be granted a downwards departure, 12pp more likely to receive a custodial sentence, and receive sentences that are 40 (33) months longer on average (at the median); (ii) Hispanics are 12pp *more* likely to be granted a downwards departure and 16pp

---

[4]In Section 5 we separately examine substantial assistance departures: these originate from the prosecution and are given on the basis of the defendant providing substantial assistance toward the prosecution of others.

[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].

more likely to receive a custodial sentence.

The next row examines whether these sentencing differentials are robust to conditioning on a rich set of covariates. Denoting the sentencing outcome for individual i of ethnicity e sentenced on day t as s_{iet}, we control for the following covariates: the demographic characteristics of the defendant described earlier (X_i) including their ethnicity (Black, Hispanic), their type of legal counsel (L_i), the offense type (OFF_{if}) where there are 31 offense types f recorded, the guideline cell they are assigned to in the pre-sentence report (G_{ig}), dummies for the Federal court district in which the case is considered (D_{id}), and year dummies $\pi_{t \in y(t)}$ where $y(t)$ is the set of days in year y. Although of course it can never be claimed that all sentencing determinants are conditioned on, a key feature of the *MCFS* data is that we can non-parametrically condition on the full set of guideline cells. This effectively proxies 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.). These factors would otherwise typically be unobservable to the econometrician.

Regressing s_{iet} on these covariates using an OLS model where the error term ε_{iet} is clustered by ethnicity-district, the next row reports the coefficients of interest $\delta_e = \{\delta_B, \delta_H\}$, measuring the partial correlation between defendant ethnicity e and sentencing outcomes.[6]

There are of course large changes in the Black and Hispanic dummy coefficients (δ_B, δ_H) as we move from the unconditional to conditional specifications along each margin. This is to be expected given defendants differ in observables by ethnicity (as Table A3a shows). However, even once we condition on rich set of covariates *including the recommended guideline cell*, there remains evidence of ethnic sentencing differentials so that relative to White defendants: (i) Blacks are significantly less likely to receive a downwards departure, significantly more likely to receive a prison sentence, and to receive significantly longer prison sentences on average and at the median; (ii) Hispanics are significantly more likely to receive a prison sentence, and they also receive significantly longer prison sentences on average and at the median. The total sentence differential of just under 4 months for Blacks (Hispanics) corresponds to 7.4% (7.8%) longer sentences for Blacks (Hispanics). The median sentence impact, that is not driven by the extensive margin, is

---

[6]To be clear, we estimate the following specification for individual i of ethnic group e sentenced on day t:

$$s_{iet} = \alpha + \sum_e \delta_e Ethnic_e + \beta X_i + \gamma L_i + \sum_f \omega_f OFF_{if} + \sum_g \gamma_g G_{ig} + \sum_d \lambda_d D_{id} + \pi_{t \in y(t)} + \epsilon_{iet}. \qquad (1)$$

s_{iet} is the sentencing outcome. We use quantile regression to estimate median impacts so as to avoid any mass point in sentence lengths given the possibility of no prison sentence being given (Table 1 shows this is the case for 22% of Whites, 10% of Blacks and 5% of Hispanics) [Powell 1986, Johnson *et al.* 2000]. The magnitude of the sentencing differentials matches a large body of historical evidence, that generally finds around a 10% Black-White differential in the probability of any custodial sentence in the Federal CJS [Bushway and Piehl 2001].

around 1.1 months for Blacks/Hispanics, suggesting that *within* the recommended guideline cell, ethnic minority defendants are sentenced for longer than Whites.

The remainder of the paper examines a candidate explanation for the Hispanic-White sentencing differences: the presence of ingroup biases that lead to 'outsiders' to US society (Hispanics) being treated differently to 'insiders' to US society (Whites). There is a vast literature examining the biological and evolutionary roots of ingroup bias, and it is often regarded as a central aspect of human behavior for individuals to aid members of a group they socially identify with, more than members of other groups they do not identify with as strongly [Tajfel *et al.* 1971]. Social psychologists have documented dimensions such as ethnicity, religiosity and political affiliation, as all being salient across contexts, in driving ingroup biases. In economics, ingroup biases have been studied in laboratory settings and show to emerge even in artificially created groups [Shayo 2009]. Field evidence on discrimination and ingroup biases in a variety of economic settings also exists [Bertrand and Duflo 2016]. Evidence of ingroup bias in judicial contexts that has been found for US State courts, as well as judicial systems outside of the US [Bushway and Piehl 2001, Shayo and Zussman 2011, Abrams *et al.* 2012, Anwar *et al.* 2012].[7] Our analysis sheds new light on whether such biases drive the Hispanic-White sentencing differentials in the high stakes and professional environment of the Federal criminal justice system.

3.2 Research Design

We consider 9-11 as an exogenously timed event that heightened the salience of insider-outsider differences and so could potentially have cued pre-existing ingroup biases. To isolate the impact the event had on sentencing outcomes in the Federal CJS, we compare outcomes between: (i) defendants who committed their last offense before 9-11 and were sentenced *before* 9-11 (a control group); (ii) to defendants who *also* committed their last offense before 9-11, but were sentenced *after* 9-11 (the treated group). We then construct a second difference in outcomes across ethnicities to estimate a difference-in-difference (DiD) impact of 9-11 on criminal sentencing. Our natural experiment (NE) sample is based on a ±180 day sentencing window around 9-11 2001, where *all*

---

[7]Bushway and Piehl [2001] use State court data from Maryland and find an impact of race on judge's decisions on sentence length (where they use guideline recommendations to instrument for unobserved case-specific characteristics). Shayo and Zussman [2011] use data from Israeli small claims courts, and exploit the random assignment of cases to Jewish or Arab judges: they find robust evidence of judicial ingroup bias. Abrams *et al.* [2012] use data from felony cases in Cook County, Illinois, and exploit the random assignment of judges to measure the between-judge variation in their treatment of race. They find statistically significant variation in how judges respond to defendant's race in incarceration rates, although not sentence lengths. They also provide evidence of ingroup bias in that Black judges have smaller Black-White sentencing gaps than White judges. On other legal actors, Anwar *et al.* [2012] show using felony trials from Florida that jury race matters for defendant outcomes: (i) juries formed from all-White jury pools convict Black defendants significantly (16pp) more often than White defendants; (ii) this gap in conviction rates is eliminated when the jury pool includes at least one Black individual.

defendants have committed their offense prior to 9-11, and hence entered the Federal CJS timeline in Figure 1, but some were sufficiently far advanced along so as to come up for sentencing pre 9-11, while others had only just entered the timeline prior to 9-11 and so ended up being sentenced post 9-11. To maintain comparability of both groups we restrict the sample further so that for those defendants sentenced before 9-11, their last offense was committed at least 180 days before 9-11.[8]

The specific advantage of exploiting 9-11 as a natural experiment is that it makes salient the kinds of ingroup bias that might drive the Hispanic-White sentencing differential in the full sample. At the same time, the research design allows us to consider Black defendants as a placebo group, whose sentencing outcomes should *not* be altered if 9-11 cues ingroup biases that are split along the lines of non-Hispanics versus Hispanics. 9-11 has been documented to have increased xeno-phobia among the US population in its immediate aftermath [Human Rights Watch 2002, Davis 2007, Woods 2011]. On individual reactions to the events of 9-11, a body of work in psychology documents how anxiety increases individual's sensitivity to risk, and that in societies with a high threat, individuals might become oversensitive to danger signals [Gadarian and Albertson 2014]. Moreover, studies in cognitive psychology suggest stress and anxiety are associated with biased information processing, where individuals tend to pay more attention to threatening information [Eysenck 1992, Yiend and Mathews 2001], and where anxiety heightens attention to threat and prioritizes the processing of threat cues [Mathews 1990].

This body of work suggests the impacts on information processing and decision making are larger among those more naturally prone to anxiety, while a key issue for our study is whether such cognitive mechanisms also impact the behavior of judges and other actors in the Federal CJS. The evidence provided in Shayo and Zussman [2011] suggests that might be so: using data from 1700 cases filed in Israeli small claims courts and exploit the random assignment of cases to Jewish or Arab judges, to provide evidence of judicial ingroup bias. By further exploiting the random timing and location of terrorist attacks in Israel, they document a short-lived difference in case outcomes depending on the ethnicity of defendants, plaintiffs and judges. To the best of our knowledge this is the only current evidence linking the salience of insider-outsider differences to ingroup biases in judicial outcomes. We extend these ideas to study the behavior of experienced legal professionals in the high stakes environment of the US Federal court system.

Table A3b shows the descriptives for the NE sample of 40,228 cases, where 32% of defendants are White, 27% are Black, and 41% are Hispanic (an ethnic composition near identical to the full sample). Moreover, there are few differences in descriptives relative to the full sample. This

---

[8]We keep cases in which: (i) guilty pleas are filed (that is so for 96% of defendants); (ii) three or fewer offenses were committed because for offenses that come up for sentencing from 01/10/2001 through to 30/09/2002, in the *MCFS* data we only observe the date of the first three offenses.

similarity across samples just reflects that defendants did not anticipate 9-11. More substantively, this implies the DiD estimate is identified from a set of criminal cases committed pre 9-11 that are representative of cases passing through the Federal CJS in other times. This is one building block for the external validity of the natural experiment.

Figure 2 provides a graphical sense of the research design by plotting histograms of the dates of sentencing and last offense for treatment and control groups, by ethnicity. Focusing first on White defendants in the top panel, the left hand histogram shows sentencing dates to be spread evenly around 9-11 as expected (with the control (treated) group entirely to the left (right) of 9-11). The right hand histogram shows the distribution of last offense dates, by treatment and control groups. By design, both groups committed their last offense before 9-11, the distribution of last offense dates in the two groups follow a similar shape, but the distribution for the treated group is right-shifted relative to the control group. The remaining panels of Figure 2 show very similar patterns for sentencing and last offense dates for treated and control groups among Black and Hispanic defendants.

Panel B of Table 1 provides descriptive evidence on sentencing outcomes, by ethnicity, for the NE sample. This replicates the findings for the full sample: pre 9-11, there are significant ethnic differentials in sentencing outcomes. The incidence of downward departures vary across ethnic groups (12.4% for Whites, 8.3% for Blacks, 28.2% for Hispanics), and for those sentenced pre 9-11, for each sentencing outcome, there is a very similar pattern of unconditional differences to Whites as in full sample. Then examining the unconditional DiD pre and post 9-11 for each ethnicity, we see that relative to Whites, Blacks and Hispanics sentenced post 9-11 are significantly *less* likely to be given a downward departure by judges: the impact on Blacks (Hispanics) is a 1.4 (4.1)pp reduction in the likelihood of a downward departure, corresponding to a 19 (17)% decline relative to the level pre 9-11. Strikingly, the unconditional Hispanic-White sentencing differential falls for average and median sentence lengths. The DiD specification we estimate is:

$$
\begin{aligned}
s_{iet} &= \alpha + \sum_e \delta_e Ethnic_e + \rho Post_t + \sum_e \phi_e (Ethnic_e \times Post_t) \\
&\quad + \beta X_i + \gamma L_i + \sum_f \omega_f OFF_{if} + \sum_g \gamma_g G_{ig} + \sum_d \lambda_d D_{id} + \epsilon_{iet},
\end{aligned}
\tag{2}
$$

where s_{iet} is the sentencing outcome for individual i of ethnic group e sentenced on day t based on a ± 180 sentencing day window around 9-11, $Post_t$ is a dummy equal to one if the defendant comes up for sentencing post 9-11, and all covariates are as defined earlier for the full sample analysis, including the full set of guideline cells that effectively proxy case-specific factors that prosecutors and legal counsel deem judges should factor into their sentencing decision. The error term ε_{iet} is

clustered by ethnicity-district. The partial correlation with ethnicity, δ_e, now captures any cross sectional differences between defendants of ethnicity tried in the Federal CJS (such as differential sorting of defendants into the Federal system), and the difference-in-difference coefficient of interest is ϕ_e. To be clear, this measures the DiD in sentencing outcomes conditional on the case reaching sentencing Stage 8. We first focus on sentencing outcomes because such decisions are cleanly attributable to Federal judges. We later study the behavior of prosecutors and legal counsel at earlier stages of the case timeline to further measure if there is differential treatment of defendants by ethnicity pre-sentencing. That moves us closer to the alternative way to measure discrimination in the CJS long debated among legal scholars, conditioning on factors that make defendants otherwise equal at the point they enter the Federal CJS in Stage 1 [Starr and Rehavi 2013].

Finally we note that 9-11 can impact sentencing outcomes for *all* defendants irrespective of their ethnicity, as measured by ρ. We refer to such impacts of 9-11 as a 'common effect'. These can arise, for example, either because anticipated changes in recidivism/detection probabilities are the same for all defendants post 9-11, or because society faces different liberty-security trade-offs post 9-11 [Davis 2007]. To measure such common effects we pool together years prior to and including 2001, thus removing any natural time effect around 9-11 each year and so isolate the common effect on sentencing outcomes for all defendants post 9-11 2001.

3.3 Identifying Assumptions and Interpreting ϕ_e

Three assumptions underpin ϕ_e identifying a causal treatment effect of ethnicity on sentencing outcomes. First, the time a defendant spends in the CJS between when their last offense is committed and when they come up for sentencing should not be impacted by 9-11. This concern is partially ameliorated by the fact that there are proscribed periods of time between each stage of the Federal CJS, and restrictions on how long some stages can take (as shown in Figure 1). The evidence in Figure 2 further points to there being no queue jumping. We address the concern more formally using survival analysis to predict the time a defendant spends in the CJS between the dates of last offense and sentencing. Second, we require there to be no ethnicity-time effects in ethnic sentencing differentials that occur naturally around 9-11 each year, say because types of criminal offense vary around the year and are correlated with defendant ethnicity. We formally assess this concern using placebo checks using data from earlier years. Finally, we require there to be no missing covariates that determine sentencing outcomes, vary across ethnic groups *and* change post 9-11 2001 (but not in placebo years). If all three assumptions hold, then on average there is no change in unobserved heterogeneity between treatment and control groups by defendant's ethnicity and ϕ_e measures the causal impact of ethnicity on sentencing differentials in the ± 180

day sentencing window around 9-11. As defendants do not anticipate 9-11, this estimate has external validity for the magnitude of ethnicity sentencing differentials in other times.

ϕ_e can reflect differences in outcomes post 9-11 driven through multiple channels. Judges might anticipate changes in behavior of defendants post 9-11, with these expectations differing across defendants by ethnicity. For example, 9-11 might have altered labor market outcomes for minorities and this can affect recidivism rates differentially across ethnic groups; alternatively, judges might anticipate post 9-11 the police will reallocate resources in a way that differentially changes future detection probabilities by ethnicity. Taken together, such channels represent different forms of statistical discrimination, where stereotyping of defendants by ethnicity can lead to differential outcomes by ethnicity post 9-11, even though all defendants in the sample were already being processed in the Federal CJS by 9-11 2001. Of course, statistical discrimination is not legally permissible because sentencing differentials cannot be justified on the basis of statistical generalizations about group traits, irrespective of whether there is an empirical foundation for this (*JEB vs. Alabama ex rel TB, 511 US 127 1994*). ϕ_e also partly captures true ethnic discrimination against ethnic group e post 9-11, and this might be especially impactful on Hispanics given the event heightened the salience of insider-outsider differences. Given these alternative interpretations of ϕ_e have different welfare implications, we later use decomposition analysis to measure how much of the unconditional DiD in sentencing outcomes can be attributed to DiD in observables or their sentencing prices across ethnicities and sentencing period, and how much remains attributable to unobservable factors.

4   Results

4.1   Judicial Sentencing Decisions

Table 2 presents estimates of (2). Focusing first on Black defendants: (i) for those sentenced pre 9-11, there are significant Black-White sentencing differentials on three out of four margins: the magnitudes of these differences are similar to those estimated in the full sample; (ii) there is no evidence of a change in Black-White sentencing gaps post 9-11 on any margin.

The pattern of results starkly differs for Hispanics, and is consistent with 9-11 exacerbating ingroup biases against Hispanics. We find that: (i) for defendants sentenced pre 9-11, Hispanic-White sentencing differentials are qualitatively and quantitatively similar to those documented for the full sample so that Hispanics receive harsher penalties in terms of the likelihood of any prison sentence, average sentence length and median sentence; (ii) for those sentenced post 9-11, Hispanic-White sentencing gaps become even more exacerbated when the salience of insider-outside

divisions is heightened: relative to Whites, the likelihood Hispanics receive a downward departure falls significantly by 3.8pp (13.5%) and the median Hispanic-White sentence differential increases by .632, suggesting that even among defendants that remain within their guideline cell, Hispanic receive significantly harsher sentences post 9-11.

To benchmark this impact we use the fact that in the Federal system, the elasticity of incarceration with respect to sentence $\simeq$ .87 [Rehavi and Starr 2014]. Given a marginal annual cost per year of imprisoning a male prisoner of $29,000$ [Congressional Research Service 2013], that implies an increase of 1329 in incarceration costs per defendant, mapping to a large increase in total costs of the Federal CJS given that 40% of all defendants are Hispanic.[9]

It is also interesting to link our findings with the established economics literature on labor market discrimination. A key insight of Gary Becker's work is that the observed racial wage gap will not reflect the average level of employer's discrimination, but rather the level of discrimination of the *marginal* employer. The reason is that minority employees can sort towards the least discriminating employer. If there is a sufficiently large share of minority workers relative to non-discriminating employers, the equilibrium wage gap reflects the tastes of the marginal employer, not the average level of discrimination in the labor market. This contrasts sharply with what we can infer in the case of criminal sentencing: as defendants cannot sort over sentencing judges, and judges cannot turn down cases they are assigned to, our estimates reflect the *average* ethnic sentencing differentials driven by judicial behavior in the Federal CJS.

The final row in Table 2 relates to the common impacts that 9-11 has on sentencing outcomes for all defendants that happened to already be in the Federal CJS on 9-11 irrespective of their ethnicity (ρ). There is evidence of common impacts on two margins: there is a 1.6% increase in any sentence being given, and average sentence lengths increase by .873 months. These effects are however identified from a simple pre-post difference and so could be driven by time effects. Hence we return to the issue below when we pool the $MCFS$ data across years and estimate placebo effects in a ± 180 day sentencing window around 9-11 for years prior to 2001.

As described in the Appendix, Tables A4 to A6 conduct a battery of checks on our core findings. These show the findings to be robust to: (i) alternative levels of clustering of the standard errors; (ii) excluding cases where perhaps because of prosecutor's decision making over the initial offense charges filed (Stage 3 in Figure 1), statutory minima or maxima bind partially over the range set

---

[9]An alternative benchmark can be based on Mueller-Smith [2016]: he uses over 2.6mn criminal cases in Texan State court data linked to individual administrative records on time in jail, unemployment insurance, public assistance benefits as well as on future criminal behavior, to estimate the total social cost generated by one year of incarceration to be between $56,000$ and $66,000$. If we apply even the lower bound estimate to our sample of defendants in the Federal CJS, this suggests the total social costs would be near double the imprisonment costs alone. As Mueller-Smith [2016] makes clear, sentencing differentials would need to have substantial deterrence effects for them to have welfare-neutral impacts.

by the guideline cell [Rehavi and Starr 2014]; (iii) estimating (2) separately for each ethnicity. Finally, we use the fact the *MCFS* data contains information on Hispanic origins and race (as described earlier, we combine both variables to construct our measure of ethnicity), to examine whether our findings pick up racial, rather than ethnic, sentencing differentials.

In Appendix Tables A7 to A10 we provide evidence in support of the underlying identifying assumptions: (i) we use survival analysis to show the time a defendant spends in the CJS between their last offense and when they come up for sentencing is not impacted by 9-11; (ii) we use data from earlier years to construct placebos 9-11 effects to check that there are no ethnicity-time effects in ethnic sentencing differentials that occur naturally around 9-11 each year.

4.2 Offense Types and Citizenship

There are two obvious reasons why Hispanic-White sentencing differentials might become exacerbated after 9-11, while Black-White differentials remain unchanged. The first relates to the fact that Hispanics are more likely to be charged with immigration offenses. If such offences are more severely punished post 9-11, ϕ_H might just pick up that Hispanics are charged with immigration offenses at a greater rate than Whites. On citizenship, Hispanics constitute the majority of non-US citizen defendants. Punishments for non-citizens, such as deportation, differ from those available for citizens/resident legal aliens, and these might become harsher for non-citizens post 9-11. If so the Hispanic-White differential would just pick up this differential selection into citizenship status. We address both issues in Table 3 by splitting the sample by offense type (drug, immigration, other), and allowing the impact of ethnicity to vary between citizens (US citizen, resident legal alien) and non-citizens (illegal aliens, non-US citizen, status unknown). 71% of defendants are thus classified as citizens; among non-citizens, 91% are Hispanic. We do so for downward departures as this is the margin along which ethnic sentencing differentials change post 9-11.

We see that for Black defendants: (i) there is no evidence of post 9-11 Black-White differentials changing for any offence type; (ii) as Black non-citizens constitute 5% of all non-citizens, those interactions are based on very small samples and so we do not focus on them (although in no case do we find evidence of sentencing differentials opening up for Black non-citizens). For Hispanics we find that: (i) the reduced likelihood of downward departures post 9-11 is largely driven by non-drug and non-immigration offenses (that constitute around 40% of all offenses); (ii) there is no evidence of within-Hispanic differences post 9-11 driven by citizenship status: the reduction in downward departures occurs for both Hispanic citizens and non-citizens alike. In short, post 9-11 Hispanic sentencing outcomes are the same irrespective of their actual citizenship status.[10]

---

[10]In line with our results, Mustard [2001] uses data on 77,000 Federal criminal cases and documents that the

4.3 External Validity

Two points are of note in relation to external validity. First, given 9-11 was unanticipated, our evidence is based on a sample of defendants and offenses that are representative of caseloads in the Federal criminal justice system more broadly. A second issue is whether the documented impacts are insightful of *pre-existing* ingroup biases against Hispanics (that might drive the sentencing differentials shown for the full sample in Table 1), or do they reflect the emergence of specific biases post 9-11? Davis [2007] provides one of the most comprehensive studies of the impacts of 9-11 on American society, and argues this included the strengthening of biases against groups considered outsiders to US society. Using data from the *National Election Survey* (NES) in 2000 that collected feeling thermometer readings where survey respondents were asked to report their attachment to various groups (as well as towards political candidates), he finds that pre 9-11 American's affection for Latinos was *lower* than for African Americans. He further notes this had also been the true in the NES data from 1976 (p203). Finally, although the picture is far from straightforward, he argues that 9-11 led to greater intolerance and uneasiness towards Latinos who became targets of anti-immigrant sentiment. At the same time, he notes that Blacks, for once, were not the targets of racial profiling and increased scrutiny (p193). These ideas map tightly to the pattern and interpretation of our core results documented in Table 2.

5 Earlier Stages of the CJS Timeline

The analysis so far has concentrated on judge's sentencing decisions. However, other legal actors are involved in earlier stages of each criminal case's timeline. Their behaviors can lead to differential treatment by ethnicity pre-sentencing, which might not necessarily be picked up in judge's sentencing outcomes. We thus extend our analysis to examine the decisions of other legal actors to understand whether they drive pre-sentencing *differential treatment* of defendants by ethnicity. We note that individual data on the ethnicity of Federal prosecutors (or legal counsel) is unavailable. However, a recent study of State prosecutors by the *Women Donors Network* (using individual data assembled by the Center for Technology and Civic Life for 2014) found that: (i) 95% of elected prosector positions are held by Whites; (ii) the majority of states have no elected Black prosecutors. Given this evidence, it is plausible that the vast majority of Federal prosecutors in the early 2000s would have been White and so the use of 9-11 as a potential trigger for raised ingroup biases against Hispanics remains valid.[11]

Hispanic-White sentence gap is generated by those convicted of drug trafficking and firearm possession/trafficking.

[11]A summary of the findings are available at http://wholeads.us/justice/wp-content/themes/phase2/pdf/key-findings.pdf (accessed May 13th 2016).

5.1 Substantial Assistance Departures

Apart from judge-initiated downward departures, another form of downward departure originates from prosecutors. They have the option to ask for a 'substantial assistance departure': this allows Federal courts to refrain from imposing a sentence within the guideline cell range on the basis of substantial assistance provided by the defendant toward the prosecution of others. The discretion to file a motion for a substantial assistance departure rests solely with the Federal prosecutor. Once such a motion is made, the sentencing judge evaluates the assistance provided and determines if such a departure is warranted. If so, the sentencing judge is granted broad discretion in determining the degree of the departure.[12]

To examine this margin of Federal prosecutor's behavior, we estimate (2) but where the outcome is whether a substantial assistance departure is granted at sentencing. The result, in Column 1 of Table 4 shows that pre 9-11, Black and Hispanics defendants *are* significantly less likely to receive substantial assistance departures than White defendants. However, post 9-11 the results indicate no evidence of any *additional* reduction in the likelihood Hispanics or Blacks being granted a substantial assistance departure.

5.2 Pre-sentence Reports

In the Federal CJS defendants must come up for sentencing precisely 75 (90) days after trial if they are held in (out of) custody. The *MCFS* data records whether a defendant is in custody after trial or not (66% of defendants are remanded in custody in the NE sample), so we can recover the precise trial date for each defendant (Stage 7 in Figure 1). This allows us to estimate the impact of 9-11 on prosecutor-legal counsel interactions that take place between trial and sentencing: this is a critical period because it is when the pre-sentence report (PSR) is drafted.

More precisely, the defendant's legal counsel provides information on the defendant's life history

---

[12]The sentencing reduction for assistance to authorities is considered independently of any reduction for acceptance of responsibility. If the prosecutor wishes to sponsor a departure from the guideline range based on the defendant's cooperation, they must make a motion under §5K1.1. Such departures are identified within the *MCFS* data. A departure from a statutory mandatory minimum penalty for cooperation requires a separate motion under 18 U.S.C. §3553(e) – these kinds of departure are not identified in the *MCFS* data. There has been some disagreement among the circuit courts as to how to determine the extent of a departure, and whether mandatory minimum sentences set limits on the extent of the departure. The USSC guidelines state that upon motion of the government stating that the defendant has provided substantial assistance in the investigation or prosecution of another person who has committed an offense, the court may depart from the guidelines. The appropriate reduction shall be determined by the court for reasons stated that may include, but are not limited to, consideration of the following: (i) the court's evaluation of the significance and usefulness of the defendant's assistance, taking into consideration the government's evaluation of the assistance rendered; (ii) the truthfulness, completeness, and reliability of any information or testimony provided by the defendant; (iii) the nature and extent of the defendant's assistance; (iv) any injury suffered, or any danger or risk of injury to the defendant or his family resulting from his assistance; (v) the timeliness of the defendant's assistance.

to the (neutral) Probation Office. The defendant is interviewed by a Probation Officer (PO), with defense counsel present. The PO collates information from this interview, forms submitted by the defense, and material provided by Federal prosecutors, to prepare a draft PSR. This is provided to the defense counsel and prosecutors 35 days before sentencing. Either party can make factual/legal objections to the draft within 10 days of receipt. A fortnight before sentencing, the final PSR is presented to the judge. This describes the defendant's background and offense (including the impact on the victim). Most importantly, it reports a determined criminal history score and the offense severity and thus calculates the recommended guideline cell and hence sentence range.

We now assess whether 9-11 impacted the suggested sentencing guideline cells differently across defendants by their ethnicity, as a result of the prosecutor-legal counsel interactions when preparing the PSR. We estimate a specification based on (2) but with two changes. First we split defendants into three groups: (i) those convicted and sentenced before 9-11 (the control group C); (ii) those convicted before 9-11, but sentenced after 9-11 ($T1$); (iii) those convicted and sentenced after 9-11 ($T2$). This three way split provides a clean comparison between the C and $T2$ group, where the latter have their PSR written *entirely after* 9-11. Second, as outcomes we consider the key recommendations from the PSR: the criminal history score, the offense severity, and the minimum sentence recommended in the implied guideline cell (hence unlike in (2), we obviously do not condition on the guideline cell).

Table 5 shows the results focusing on the clean comparison between the C and $T2$ group of defendants: we find no evidence of differential impacts post 9-11 on either Hispanic nor Black defendants for five out of six dimensions of the PSR. Hence, prosecutor-legal interactions at the PSR stage between trial and sentencing are *not* a source of differential treatment of defendants by ethnicity post 9-11 when insider-outsider differences are most salient.[13]

We do however find evidence of a *common* impact of having the PSR written after 9-11: significantly higher offense severity scores are recommended, and the consequent minimum sentence in the guideline cell significantly rises by 2.6 months. Furthermore, the earlier results on the behavior of judges *conditional* on these recommendations suggest judge's *reinforce* the harsher treatment defendants experience post 9-11 (rather than judge's acting to offset harsher PSR's being drafted post 9-11): Table 2 showed these common impacts post 9-11 (ρ) to be a 1.6% increase in any sentence being given, and an average sentence length increase of .873 months. The interplay between prosecutor and judges decisions is a novel insight that our data and research design allows to be documented.

---

[13]For those defendants in $T1$ we also find no impacts on these PSR outcomes: these are harder to interpret because these PSRs will be drafted both pre- and post 9-11.

5.3 Prosecutor's Initial Offence Charges

Earlier in the timeline of Federal criminal cases, a critical decision prosecutor's have discretion over is the initial offense charges filed (Stage 3 in Figure 1). In the Federal criminal code, definitions of crimes often overlap, providing prosecutors discretion over the initial offense to charge. These initial offense charges are crucial because: (i) they determine if statutory minima/maxima sentences bind and take precedence over the guideline cell sentence range; (ii) they determine outside options in plea bargaining, and in 80% of cases remain unchanged [Rehavi and Starr 2014].[14]

Rehavi and Starr [2014] establish using similar linked administrative data that prosecutor's initial offense charges account for half the Black-White sentencing gap in the later period 2006-8, after sentencing guidelines had been abolished and judges are not required to issue sentences within the guidelines. We first establish whether their findings replicate in the full sample period of our study 1998-2003, when sentencing guidelines are always in place. To do so we estimate a specification analogous to (2) but where the outcomes considered are: (i) whether the defendant receives an initial charge with a non-zero statutory minimum sentence; (ii) the length of statutory minimum sentence associated with their initial offense. In this specification we do not condition on offense type, or guideline cell (the former because the offense charge might go across offense type boundaries, and the latter because it is determined later in the timeline). This analysis therefore more closely measures differential treatment by ethnicity conditional on factors that make defendants otherwise equal at the point they *enter* the Federal CJS.

The results are in Columns 2 and 3 of Table 4. Precisely in line with Rehavi and Starr [2014] we find that in our full sample of 230,000 criminal cases from October 1998 through to September 2003, prosecutors set initial offense charges significantly more harshly for Black and Hispanic defendants, both on the extensive margin of a non-statutory minimum sentence being set and the actual length of the statutory minimum. This confirms that when sentencing guidelines are place, this margin is a key one along which prosecutor's actions determine ethnic sentencing differentials: this is exactly the point made by Rehavi and Starr [2014]. The magnitude of the effect is such that Blacks receive charges carrying minimum sentences that are 21 months longer than Whites, corresponding to 52% of the raw Black-White sentencing gap in Table 1.

---

[14]Many forms of statutory minima exist and can have precedence over the minimum from the guideline cell. In 15.8% (3.6%) of cases the statutory minimum is above (below) the guideline minimum (maximum). Rehavi and Starr [2014] provide an example of how prosecutor's need to assess the strength of evidence, and characterization of ambiguous facts determine initial offense charges. This relates to the use of firearms in a burglary. If a gun is found in the car that transported a defendant to a burglary, the prosecutor must decide whether to allege the burglary legally qualified as a "crime of violence", that the gun qualified as a firearm, and that the defendant "carried" it "during and in relation to" the burglary. All these factors are necessary to trigger a five year mandatory sentence, and would run consecutively to the burglary sentence. Rehavi and Starr [2014] point out a lenient prosecutor might choose to "swallow the gun" and just charge the burglary.

We next examine whether the events of 9-11, that heightened the salience of insider-outsider differences, lead to these ethnic differentials being widened for Hispanics (but not for Blacks). To pinpoint the ingroup biases of prosecutors, we consider a narrow window covering a cohort of 3600 defendants *all* of whom entered the Federal system pre 9-11 but had their initial offense charges filed either side of 9-11. Taking the date of last offense as a proxy for time of entry into the Federal CJS (Stage 1), we exploit the fact that the system requires defendants in (out of) custody to have their initial offense charges brought within 14 (21) days. This allows us to define two groups of defendant: (i) those whose last offense was committed 29 to 42 (43 to 63) days before 9-11 (depending on whether they are in custody or not) and so whose initial offense charge was determined prior to 9-11 (a control group); (ii) those whose last offense was committed 14 (21) days before 9-11 until the day before 9-11 and so their initial offense change would have been determined just after 9-11 (a treated group). We then estimate a specification analogous to (2) but where the outcomes considered are: (i) whether the defendant receives an initial charge with a non-zero statutory minimum sentence; (ii) the length of statutory minimum sentence associated with their initial offense. As before we do not condition on offense type or guideline cell.[15]

The results are shown in Table 6 and confirm that: (i) Hispanic defendants initially charged post 9-11 are 7.5pp more likely to receive an initial offense that carries a statutory minimum corresponding to a 100% increase over the pre 9-11 period; (ii) their statutory minimum sentence is 10.7 months longer; (iii) there is no evidence that 9-11 impacts prosecutor's initial offense charges filed against Black defendants along either margin ($\widehat{\phi}_B = 0$ in Columns 1 and 2).

The remaining Columns then trace through the sentencing impacts on the same cohort of defendants. We thus compare defendants who all come up for sentencing post 9-11, but vary in whether their initial offense charge was filed pre or post 9-11. We do not condition on offense type or guideline cell as both are determined by prosecutor's choice of initial offense charges. We see that for Hispanics who were initially charged just after 9-11, the higher statutory minimum associated with their charge translates into significantly longer sentences of 9.3 months. The differential pre-sentencing treatment of this cohort of defendants represents additional large additional incarceration costs per defendant of ingroup biases that we have not so far measured. The earlier costs were associated with the cohort that come up for sentencing around 9-11 (Table 2) whereas these results imply continuing *longer run* costs of ingroup biases that relate to the cohort of Hispanic defendants initially charged around 9-11, and come up for sentencing well after 9-11.

Finally, in Columns 4 and 5 we control for offense type and guideline cell and find no difference in judicial sentencing outcomes for this cohort in sentence length or downward departures. What

---

[15]We remove those whose last offense was committed 15 to 28 (22 to 42) days before 9-11 to avoid mis-classifying these individuals into which side of 9-11 their initial offense charge was filed.

these specifications imply is that judges do not offset the differential behavior of prosecutor's towards Hispanics around 9-11 with regards to initial offense charges. This further interplay between prosecutor and judges decisions is a novel insight that our data and research design allows to be documented.

5.4 Decomposition Analysis

The analysis has highlighted two cohorts of Hispanic defendant for whom 9-11 led to greater sentencing disparities relative to Whites: (i) for those cohorts that come up for judicial sentencing just after 9-11, Hispanics are significantly less likely to receive downward departures (Table 2); (ii) for those cohorts for whom prosecutors set initial offence charges just after 9-11, Hispanics receive charges associated with significantly longer statutory minimum sentence lengths (Table 6). To rule out some potential drivers of these differentials, we use a Juhn-Murphy-Pierce [1993] decomposition to split the raw DiD in sentencing outcomes into those attributable to: (i) changes in the observable characteristics of defendants; (ii) changes in the returns to these observables (or changes in the sentence 'price' of observables); (iii) changes in unobservables. The JMP decomposition is implemented by first considering the following sentencing equation for White defendant i sentenced in period T: $s_{iT} = X'_{iT}\beta^W_T + \sigma^W_T \theta_{iT}$, where β^W_T are sentence prices for Whites, θ_{iT} is a standardized residual capturing unobserved determinants of White sentences, and σ^W_T is the standard deviation of this residual for Whites in period T. The Hispanic-White sentencing differential in period T is then, $\Delta s_T = s^H_T - s^W_T = \Delta X_T \beta^W_T + \sigma^W_T \Delta \theta_T$. Given our DiD research design we take a *second* difference over time periods, considering how the ethnic sentencing gap changed pre- to post 9-11 ($T = 0$ to $T = 1$):

$$\Delta s_1 - \Delta s_0 = (\Delta X_1 - \Delta X_0)\beta^W_0 + \Delta X_1(\beta^W_1 - \beta^W_0) + (\Delta \theta_1 - \Delta \theta_0)\sigma^W_0 + \Delta \theta_1(\sigma^W_1 - \sigma^W_0). \quad (3)$$

The $(\Delta X_1 - \Delta X_0)\beta^W_0$ component, or X-effect, measures the contribution to the DiD in sentencing gaps of observables. The $\Delta X_1(\beta^W_1 - \beta^W_0)$ component, or β-effect, measures changes in sentencing prices pre- and post 9-11 for all these observables. For example, some offense types, such as those related to immigration, might be punished more harshly post 9-11 due to changes in expectation over defendant's future recidivism or detection probability. These impacts also capture changes in the sentencing price of being in each recommended guideline cell, G_{ig}. These recommendations embody case-specific information that prosecutors and legal counsel deem relevant for judge's sentencing decisions, such as whether a firearm is used, or for drug offenses, the quality of drugs in the case etc.

While it is well understood that such decompositions in no way represent formal tests for statistical discrimination [Charles and Guryan 2011], in our setting the usual concerns related to decomposition analysis for studying discrimination are partly ameliorated because: (i) the DiD set-up reduces concerns over lack of common support in the cross-section of covariates across ethnicities; (ii) the inclusion of guideline cell dummies allows us to capture many more case-specific factors driving sentencing outcomes than would normally be observed by the econometrician. With these issues in mind, the combined X- and β-effects can potentially encapsulate multiple channels through which statistical discrimination can operate, or channels through which post 9-11 sentencing might justifiably respond.

The $(\Delta\theta_1 - \Delta\theta_0)\sigma_0^W$ component, or θ-effect, measures the change in Hispanic's position within the White residual sentencing distribution (measured at $T = 0$). Shifts in discrimination against Hispanics post 9-11 would lead to an increase in Hispanic's average position in the White residual distribution. Finally, the $\Delta\theta_1(\sigma_1^W - \sigma_0^W)$ component, or σ-effect, measures changes in the spread of the White sentencing residual from pre- to post 9-11, holding fixed the post 9-11 ethnic residual gap $\Delta\theta_1$. The θ-effect and σ-effect reflect both discrimination and unobservable offense and defendant characteristics. *A priori* we might expect the θ-effect to predominantly reflect shifts in ethnic discrimination because it represents changes in the *position* of Hispanics in the White sentencing residual distribution, while the σ-effect captures changes in the *spread* of this residual, that is less clear would be driven by ethnic discrimination.

Table 7 shows the JMP decomposition for Hispanic-White sentencing gaps for the two cohorts identified above. On judge's sentencing decisions, the decomposition for downward departures (the margin along which ethnic sentencing differentials change post 9-11) is based on a LPM.[16]

Column 1 shows that: (i) only 7% is attributable to observables (Row 4: X-effect + β-effect); (ii) 93% of the Hispanic-White differential is due to unobservables (Row 5: θ-effect + σ-effect); (iii) among the unobservable components, the θ-effect is by far the more important driver of the unconditional DiD in downward departures, namely change in Hispanics' position within the White residual sentencing distribution (measured at $T = 0$) (Row 8); (iv) there is not much evidence of a change in the spread of the White residual: the σ-effect is only .006 (Row 9).[17]

Column 2 focuses on the cohort of defendants impacted by prosecutor's initial offence charges.

---

[16]To check the validity of basing the JMP decomposition off a linear probability, we have also conducted cross-sectional decompositions in the pre- and post 9-11 periods separately using both a Blinder-Oaxaca decomposition and the Fairlie [2005] extension of such decompositions to non-linear models. Constructing the implied difference-in-difference decomposition from either the Blinder-Oaxaca or the Fairlie approach, generates very similar conclusions as the JMP decomposition based on the LPM.

[17]In contrast, for other judicial sentencing margins (any sentence, sentence), the majority of the Hispanic-White differential *is* explained by observables (X-effect + β-effect). This matches with the earlier evidence where the sentencing differential post 9-11 was found to only become exacerbated along the downward departures margin.

For this continuous outcome the application of the JMP decomposition is straightforward, and in line with the earlier results we do not control for offence type or guideline cell in the set of X's. Column 2 shows that: (i) based on observables, the Hispanic-White gap would be predicted to *fall* post 9-11 not rise (Row 4: X-effect + β-effect); (ii) unobservable factors entirely drive the Hispanic-White differential and among the unobservable components, the θ-effect is by far the more important driver of the DiD in statutory minimum sentence lengths.

Taken together the results suggest that for both sets cohorts of Hispanic defendant for whom 9-11 led to greater sentencing disparities relative to Whites, neither disparity is easily explained by changes in observables or the sentencing prices of observables.

6 Judges and Ingroup Biases

Judges ultimately make sentencing decisions and so can choose to offset, reinforce or ignore the differential treatment of offenders by prosecutors etc. arising earlier in the timeline of criminal cases. In this Section we therefore analyze how judge characteristics correlate to our measure of Hispanic-White sentencing differentials to shed light on whether ingroup biases of judges might be the origin of these differentials. The administrative data contains no information on judges, and there is no way to link judge and defendant identifiers for Federal criminal cases. To make progress we have hand-coded the characteristics of Federal judge's by district, sourced from the *Biographical Directory of Federal Judges*. This details the ethnicity, gender, and seniority of judges in 90 districts d, as well as whether they were appointed under a Democrat or Republican President. As described further in the Appendix, we thus construct judge characteristics at the district level ($\mathbf{J}_d$).

We then proceed in two steps. First, we estimate (2) allowing for a full set of interactions between each Federal district d and ($Hispanic_e \times Post_t$) to estimate the coefficient of interest: $\phi_{H,d}$. We do so for the likelihood a downward departure is given as this is the margin along which ethnic sentencing differentials further open up post 9-11. Second, we regress $\widehat{\phi}_{H,d}$ against $\mathbf{J}_d$, where observations are weighted by the share of defendants in district d in the NE sample that are Hispanic, and robust standard errors are reported. Observations are weighted because the underlying regression from which each $\widehat{\phi}_{H,d}$ is estimated is based on individual observations, and this number varies by district. In contrast to Federal prosecutors, there are a substantial share of judges from minority backgrounds. The weighted mean share of Hispanic (Black) judges in a district is 14% (7%); 17% of judges are women, 28% are of senior status, and 48% are appointed by Democrat Presidents. As there are only on average 7.5 judges per district, small changes in

the composition of judges can significantly alter a defendant's probability to be sentenced by a minority judge.[18]

Table 8 shows the second stage results. In Column 1 we only control for judge ethnicities: we see that in districts where there are a higher proportion of Hispanic judges, the Hispanic-White sentencing differential, as measured by $\widehat{\phi}_{H,d}$, is significantly smaller. This is in line with judges displaying ingroup bias towards defendants along the lines of insider-outsider divisions [Schanzenbach 2005, Abrams *et al.* 2012].

Column 2 shows this finding to be robust to controlling for the seniority, gender and appointment characteristics of Federal district judges. Hence the Hispanic ethnicity of judges is not merely picking up them being Democrat appointees, and consistent with the evidence in Schanzenbach [2005], the presence of Democratic appointed judges has an independent correlation with Hispanic-White sentencing differentials, all else equal. Column 3 additionally controls for the population shares of different ethnic groups in the state in which the district belongs, as well the change (1990 to 2000) in the proportion of the population from each ethnic group in the state. Doing so *increases* the coefficient on the district proportion of Hispanic judges from .174 to .331 (where both are significant at conventional levels) and this partial correlation becomes more precisely estimated. Hence the district proportion of Hispanic judges does not appear to be proxying for population characteristics of where the Federal criminal case is heard.

Figure 3 provides a sense of the spatial pattern of sentencing differentials. Panel A shows the population of Hispanics by State in 2000: as expected, Southern and Western States are those in which the highest shares of the population are Hispanic (the map for the change in population shares is very similar). Panel B then plots $\widehat{\phi}_{H,d}$ for each district d: this highlights that ingroup bias against Hispanics is generally more pronounced in districts where a smaller share of the population is Hispanic. This pattern of results is very much in line with the contact hypothesis, that states that interpersonal contact is an effective ways to reduce prejudice between majority and minority group members [Allport 1954].

There is also evidence across specifications that senior judges, female judges and those appointed under Democrat Presidents are also less likely to discriminate against Hispanics. As such this offers a suggestion that returns to experience might also help ameliorate ethnic discrimination against Hispanic defendants. To more easily make comparisons across covariates, Column 4 standardizes each covariate and reports effect size estimates of each partial correlation. We see that a one standard deviation in the proportion of judges in the district of Hispanic origin increases $\widehat{\phi}_{H,d}$

---

[18]Senior judges are partially retired and have greater discretion over their caseload. Schanzenbach [2005] provides evidence that the absolute number of Hispanic Federal judges has been relatively constant over the period from 1990 to 2002; the rises in the number of Black and female judges are considerably more pronounced.

by 2.1pp. This effect size is comparable to the implied impact on the Hispanic-White sentencing differential of a one standard increase in the share of senior or female judges. The effect size is also comparable in absolute magnitude to the average effect across all districts, documented in Table 4 that post 9-11, Hispanic defendants are 3.8pp less likely to receive a downward departure.

To examine the external validity of these correlations outside of the window around 9-11, the next Column repeats the exercise but first estimates $\widehat{\delta}_{H,d}$ from the full sample of $230,000$ Federal criminal cases from October 1998 to September 2003. We continue to report all coefficients as effect sizes to aid comparability. Strikingly, in the full sample we also see evidence of ingroup bias: a one standard deviation in the proportion of district judges of Hispanic origin increases the Hispanic-White sentencing differential for downward departures, $\widehat{\delta}_{H,d}$, by .028, that is actually slightly larger than the effect size estimate based on the NE sample estimates. As a final step in examining whether ingroup biases drive the behavior of judges, the final Column uses (1) to first estimate $\widehat{\delta}_{B,d}$ from a full set of interactions between each Federal district d and $Black_e$. We then see that: (i) Black defendants are punished significantly more harshly by Hispanic judges (in line with insider-outsider divisions being drawn between Hispanics and non-Hispanics); (ii) Black Federal judges appear not to treat Black defendants differently from White defendants (again in line with insider-outsider divisions not being drawn *within* non-Hispanics).[19]

7   Conclusions

A large body of literature across disciplines documents that for similar offences, Blacks and Hispanics face a higher probability of arrest, conviction and harsher penalties conditional on conviction. If historic trends continue, then among the 2001 birth-cohort, one in three Black men and one in six Hispanic men can expect to spend time in prison during their lives [CEA 2016]. The central challenge lies in understanding whether such differential outcomes in the criminal justice system by ethnicity are driven by unobserved heterogeneity across defendants, correlated to their ethnicity, or whether they reflect true discrimination. The primary reason research on sentencing differentials has been deadlocked is because the origins of such unobserved heterogeneity can stem from so many sources such as: (i) the characteristics of those arrested by the police and assigned to be tried in the CJS; (ii) the behavior of judges, prosecutors, legal counsel and defendants, during the various stages of the CJS. In this paper we present a difference-in-difference research design that addresses the first concern, and we utilize linked administrative data to tackle the second issue.

---

[19]On other evidence consistent with such Black-Hispanic divisions, Holzer and Ihlanfeldt [1998] provide evidence from a survey of employers that shows Blacks are a less likely to get hired in firms with majority Hispanic customers than in firms with majority White customers.

We do so in the high stakes environment of the Federal criminal justice system, where decisions are made by professional and experienced judges, prosecutors and legal counsel, and the universe of criminal offenses and district courts can be studied.

The key contribution of our analysis has been to combine linked administrative data with a research design to provide insights into the magnitude, channels and potential origins of Hispanic-White sentencing differentials in the Federal criminal justice system. This is important given Hispanics are: (i) the modal defendant in the Federal CJS; (ii) the group whose incarceration rate is growing fastest; (iii) a relatively understudied minority group in the economics of discrimination literature [Charles and Guryan 2011], despite Hispanics being ever more prominent in the political, legal and cultural life of America. We build an evidence base that is consistent with these differentials being driven by the ingroup biases of judges and prosecutors, whereby 'outsiders' (Hispanics) are treated differently to 'insiders' (Whites). Going beyond the existing literature we document evidence of such ingroup biases driving the behavior of Federal judges at sentencing, and the behavior of prosecutor's when setting initial offense charges.

Our results offer the suggestion that increasing accountability at these stages of Federal criminal cases, or appointing more Hispanic judges to Federal district courts might go some way towards reducing Hispanic-White sentencing differentials in the criminal justice system. By documenting the potential origins of Hispanic-White sentencing differentials, we add to Shayo and Zussman [2011] as providing evidence linking the salience of insider-outsider differences to ingroup biases in judicial outcomes. More broadly, our analysis helps address an appeal made in recent overviews of the economics of discrimination literature on the need to better bridge to the psychology literature on the origins of discriminatory behavior [Charles and Guryan 2011, Bertrand and Duflo 2016].

A natural next step would be to use such linked administrative data to push forward the research frontier on the origins of Black-White sentencing differentials. There is of course a vast literature in social psychology suggesting stereotyping of Blacks might lie at the root of such differences; laboratory experiments provide foundational evidence for this based on visual processing [Eberhardt *et al.* 2004], and recent field experiments also highlight the role that limited attention might play in driving discrimination [Bartos *et al.* 2016]. The challenge lies in developing credible research designs in the context of the criminal justice system that cause the strength of such factors underpinning the origins of discrimination to vary across time or space in a manner orthogonal to other characteristics of criminal cases. Given the social, legal and economic consequences of how the criminal justice system is differentially experienced by individuals of different ethnicities, we hope our findings here on the origins of Hispanic-White sentencing gaps encourage others to also take up this challenge.

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] AIZER.A AND J.DOYLE (2015) "Juvenile Incarceration, Human Capital and Future Crime: Evidence from Randomly-assigned Judges," forthcoming, *Quarterly Journal of Economics.*

[3] ALESINA.A AND E.LA FERRARA (2015) "A Test of Racial Bias in Capital Sentencing," *American Economic Review* 104: 3397-433.

[4] ALLPORT.G.W (1954) *The Nature of Prejudice*, Cambridge, MA: Perseus Books.

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

[6] BARTOS.V, M.BAUER, J.CHYTILOVA AND F.MATEJKA (2016) "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition," *American Economic Review* 106: 1437-75.

[7] BAYER.P, R.HJALMARSSON AND D.POZEN (2009) "Building Criminal Capital Behind Bars: Peer Effects in Juvenile Corrections," *Quarterly Journal of Economics* 124: 105-47.

[8] BERTRAND.M AND E.DUFLO (2016) "Field Experiments on Discrimination,", forthcoming in A.Banerjee and E.Duflo (eds.) *Handbook of Field Experiments.*

[9] BUSHWAY.S.D AND A.M.PIEHL (2001) "Judging Judicial Discretion: Legal Factors and Racial Discrimination in Sentencing," *Law and Society Review* 35: 733-64.

[10] CEA (2016) *Economic Perspectives on Incarceration and the Criminal Justice System*, Washington DC.: Council of Economic Advisors Report.

[11] CHARLES.K.K AND J.GURYAN (2011) "Studying Discrimination: Fundamental Challenges and Recent Progress," *Annual Review of Economics* 3: 479-511.

[12] CONGRESSIONAL RESEARCH SERVICE (2013) *The Federal Prison Population Buildup: Overview, Policy Changes, Issues and Options*, Report 7-5700, Washington DC: CRS.

[13] DAVIS.D (2007) *Negative Liberty: Public Opinion and the Terrorist Attacks on America*, Russell Sage Foundation.

[14] DI TELLA.R AND E.SCHARGRODSKY (2013) "Criminal Recidivism after Prison and Electronic Monitoring," *Journal of Political Economy* 121: 28-73.

[15] EBERHARDT.J.L, P.A.GOFF, V.J.PRUDIE AND P.G.DAVIES (2004) "Seeing Black: Race, Crime and Visual Processing," *Journal of Personality and Social Psychology* 87: 876-93.

[16] EYSENCK.M.W (1992) *Anxiety: The Cognitive Perspective*, Mahwah, NJ: Lawrence Erlbaum.

[17] FAIRLIE.R.W (2005) "An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models," *Journal of Economic and Social Measurement* 30: 305-16.

[18] GADARIAN.S.K AND B.ALBERTSON (2014) "Anxiety, Immigration, and the Search for Information," *Political Psychology* 35: 133-64.

[19] HOLZER.H.J AND K.R.IHLANFELDT (1998) "Customer Discrimination and Employment Outcomes for Minority Workers," *Quarterly Journal of Economics* 113: 835-67.

[20] HUMAN RIGHTS WATCH (2002) We Are Not the Enemy: Hate Crimes Against Arabs, Muslims, and Those Perceived to be Arab or Muslim after September 11, Human Rights Watch 6.

[21] JOHNSON.W, Y.KITAMURA AND D.NEAL (2000) "Evaluating a Simple Method for Estimating Black-White Gaps in Median Wages," *American Economic Review* 90: 339-43.

[22] JUHN.C, K.M.MURPHY AND B.PIERCE (1993) "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy* 101: 410-42.

[23] KLEPPER.S, D.NAGIN AND L-J.TIERNEY (1983) "Discrimination in the Criminal Justice System: A Critical Appraisal of the Literature," in Research on Sentencing: The Search for Reform, A.Blumstein, J.Cohen, S.E.Martin and M.H.Tonry (eds.) Vol. 2. Washington DC: National Academy Press.

[24] KLING.J (2006) "Incarceration Length, Employment, and Earnings," *American Economic Review* 96: 863-76.

[25] MATHEWS.A (1990) "Why Worry? The Cognitive Function of Anxiety," *Behaviour Research and Therapy* 31: 57-62.

[26] MUELLER-SMITH.M (2016) The Criminal and Labor Market Impacts of Incarceration, mimeo, University of Michigan.

[27] 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.

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

[29] POWELL.J (2006) "Censored Regression Quantiles," *Journal of Econometrics* 32: 143-55.

[30] REHAVI.M.M AND S.B.STARR (2014) "Racial Disparity in Federal Criminal Sentences," *Journal of Political Economy* 122: 1320-54.

[31] SCHANZENBACH.M (2005) "Racial and Sex Disparities in Prison Sentences: The Effect of District-Level Judicial Demographics," *Journal of Legal Studies* 34: 57-92.

[32] SHAYO.M (2009) "A Model of Social Identity with an Application to Political Economy: Nation, Class and Redistribution," *American Political Science Review* 103: 147-74.

[33] SHAYO.M AND A.ZUSSMAN (2011) "Judicial Ingroup Bias in the Shadow of Terrorism," *Quarterly Journal of Economics* 126: 1447-84.

[34] STARR.S.B AND M.M.REHAVI (2013) "Mandatory Sentencing and Racial Disparity: Assessing the Role of Prosecutors and the Effects of *Booker*," *Yale Law Journal* 123: 2-80.

[35] TAJFEL.H, M.G.BILLIG, R.P.BUNDY AND C.FLAMENT (1971) "Social Categorization and Intergroup Behavior," *European Journal of Social Psychology* 1: 149-78.

[36] 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.].

[37] WOODS.J (2011) "The 9/11 Effect: Toward a Social Science of the Terrorist Threat," *Social Science Journal* 48: 213-33.

[38] YIEND.J AND A.MATHEWS (2001) "Anxiety and Attention to Threatening Pictures," *Quarterly Journal of Experimental Psychology* 54: 665-81.

A Appendix

A.1 Data Sources

The data used were obtained from the Inter-university Consortium for Political and Social Research and are part of the *MFCS* series, derived from cases received by the USSC. As described in Rehavi and Starr [2014], the four linked data sets are: (i) US Marshals Service (USMS) data, that covers the arrest/offense stage (Stage 0) and includes all persons arrested by Federal law enforcement agencies, persons arrested by local officials and then transferred to Federal custody, and persons who avoid arrest by self-surrendering; (ii) Executive Office for US Attorneys (EOUSA) data, covering initial appearance through to arraignment (Stages 1-3): these data come from the internal case database used by Federal prosecutors, and covers every case in which any prosecutor at a US Attorney's office opens a file; (iii) Administrative Office of the US Courts (AOUSC) data, covering initial district court appearances through to trial (Stages 4-7): these originate from Federal Courts and contain data on all criminal cases heard by Federal district judges, and any non-petty charge handled by a Federal magistrate judge; (iv) US Sentencing Commission (USSC) data, covering the sentencing Stage 8: this data set collects information on any case that results in conviction and sentencing for a non-petty offense. These data are collected by the Bureau of Justice Statistics. We drop 4 out of 94 districts: Guam, Puerto Rico, N.Mariana Island and the Virgin Islands.

We focus on male defendants that come up for sentencing from October 1998 to September 2003. We focus on this period because: (i) before October 1998 the data is less detailed; (ii) from October 2003 sentencing guidelines began to be reformed.[20]

The types of downward departure listed in the USSC sentencing guidelines and coded in the data are: (i) encouraged departure factors (those that take into factors such as coercion or duress, diminished capacity, or aberrant behavior of nonviolent offenders); (ii) discouraged departure factors (such as age, physical condition, family responsibilities, or prior good works); (iii) unmentioned factors that were not adequately considered by the guidelines (such as extraordinary rehabilitation after the offense but prior to sentencing). The last group are the most frequently cited type of downward departures (82% of the total), and this is so for all ethnicities.

The data for judicial characteristics are sourced from the *Biographical Directory of Federal Judges*. To select the relevant judges to construct the district-level judge characteristics , we used the data on commission and termination dates for each judge in the database, we restrict the sample to judges commissioned before the end of the natural experiment sample and those who

---

[20]More information on the data series can be found at,
http://www.icpsr.umich.edu/icpsrweb/NACJD/series/00083/studies?archive=NACJD&sortBy=7 (accessed 14th April 2016).

terminated the bench after the beginning of the sample. We perform an analogous sample cut of judges relevant for the full sample specifications.

A.2 Robustness Checks

The main specifications cluster standard errors by ethnicity-district and so focus on geographically based unobservables that might be correlated by ethnicity for sentencing outcomes. The alternative level of clustering we therefore consider is at the level of week of sentencing x ethnicity, so placing more emphasis on *time-related* unobservables being correlated by ethnicity for sentencing outcomes. The resulting standard errors are near identical to those in Table 3 in most cases (Table A4, Columns 1a-1d).

The second check excludes cases where statutory minima or maxima bind partially over the range set by the guideline cell [Rehavi and Starr 2014]. This occurs in 19% of cases, but the estimated effects follow a similar pattern to those estimated on the NE sample (Table A4, Columns 1a-1d). In Section 5 we explicitly examine whether post 9-11, prosecutor's change their decisions over the initial offense charges to file at Stage 3 post 9-11 differentially across ethnicities. Table A5 shows the core results to be robust to estimating (2) separately for each ethnicity: the signs, significance and magnitude of estimates matches closely the pooled specification, with there remaining an implied DiD penalty of a 3.4pp reduction in the likelihood Hispanic defendants are granted downward departures if sentenced post 9-11.

On racial sentencing differentials, Table A6 shows the results, where we estimate a specification analogous to (2) but allow the post 9-11 impacts to vary by race, using the full set of race classifications in the *MCFS* data. To establish the link between this split and what we have previously used, it is important to note that defendants we coded as Hispanics are, in this specification, spread over those coded as white- or black-race, but with 92% of them being white-race. Strikingly, we find no evidence of racial sentencing differentials opening up post 9-11, relative to white-race defendants. Our main results thus point to ethnic, rather than racial sentencing differentials. The main document Hispanic-White ethnic sentencing differential is simply masked in this specification within the white-race impacts.

A.3 Evidence in Support of the Identifying Assumptions

To underpin a casual interpretation of the results, we first examine the identifying assumption that the time a defendant spends in the Federal CJS between when they commit their last offense and when they come up for sentencing is not impacted by 9-11. Table A7 first addresses this concern by extending specification (2) to additionally control for the defendant's time in the CJS

using two approaches: (i) include a series of dummies grouping the time between the last offense and sentence date; (ii) including a series of dummies grouping the last offense date. As shown in Table A7, the earlier results are robust to using either approach (which is not too surprising given the descriptive evidence in Figure 2). A direct test of this identifying assumption is provided in Table A8 where we use OLS and survival models to estimate the time between last offense and sentencing date for each defendant, and then test whether this changes significantly, by ethnicity, post 9-11. The survival models used are the nonparametric Cox and the log logistic model because it allows for a frailty parameter. Across specifications we find no robust evidence of a change in time defendants spend in the Federal CJS post 9-11, by ethnicity (Columns 1a-1c). Nor do we find any evidence of longer processing times for all defendants (the coefficient on $Post_t$ is not different from zero). These findings also hold just for specific offense types (Columns 2a-4c).

The second identifying assumption is that there are no ethnicity-time effects in ethnic sentencing differentials that naturally occur around 9-11 each year. We use the $MCFS$ data on cases from earlier years (1999 onwards) to estimate placebo 9-11 impacts by ethnicity.[21] The results are shown in Table A9 and confirm that: (i) there are no natural ethnicity-time effects around 9-11 along any sentencing margin; (ii) the earlier documented impacts for Hispanics only occur post 9-11 in 2001. At the foot of each Column we provide an estimate of the common impact effect for each sentencing margin, that takes into account any ethnic specific time trends around 9-11 each year. We see the likelihood of receiving a prison sentence increases by 2.3% for all defendants. The sentence length effect remains positive and of similar magnitude to that estimated in Table 2, but is imprecisely estimated.[22]

---

[21]The sample of criminal cases used are those $114,642$ cases for which sentencing occurs within a 6-month window of 9-11 in years 1998 to 2003 and: (i) if sentenced after 9-11, the last offense was committed prior to 9-11 each year; (ii) if sentenced before 9-11, the last offense was committed up to 6-months prior to 9-11 that year.

[22]A natural candidate to explain this common effect is the introduction of the Patriot Act on the 26th of October 2001. This made important changes to how certain Federal offenses were treated (especially those related to immigration and money laundering), and might also have reflected different trade-offs and permanently altered objectives of the Federal CJS post 9-11. To examine whether the Patriot Act relates to the common effect on the any sentence margin, Table A10 presents estimates a modified specification based on (2) but that further splits the treated group of defendants into those: (i) sentenced post 9-11 and before the introduction of the Patriot Act; (ii) sentenced post 9-11 and before the introduction of the Patriot Act. We see the common impact on the margin of any prison sentence being imposed is driven entirely by those cases that are up for sentencing after the Patriot Act is passed (Column 1): the magnitude of the common effect is .017 (compared to the .023 common effect estimated earlier in Table A9). The remaining Columns show that: (i) these common impacts apply to both Patriot Act offenses (immigration and money laundering) and non Patriot Act offenses, although the magnitude is far larger for Patriot Act offenses as expected (.080 versus .015).

Table 1: Sentencing Outcomes by Ethnicity

Standard errors in parentheses, clustered by ethnicity-district

	White				Black				Hispanic			
	(1) Downwards Departure	(2) Any Sentence	(3) Sentence	(4) Median	(5) Downwards Departure	(6) Any Sentence	(7) Sentence	(8) Median	(9) Downwards Departure	(10) Any Sentence	(11) Sentence	(12) Median
A. Full Sample	.122	.784	42.6	24	.074	.903	82.9	57	.244	.948	41.8	27
Unconditional difference to Whites					-.047**	.120***	40.3***	33.0***	.122**	.165***	-.833	3.00
					(.019)	(.012)	(3.44)	(3.27)	(.060)	(.011)	(4.05)	(6.66)
Conditional difference to Whites					-.010*	.028***	3.97***	1.05***	.002	.059***	3.69***	1.32***
					(.006)	(.004)	(.468)	(.216)	(.015)	(.006)	(.487)	(.258)
B. Natural Experiment Sample												
Last offense pre 9-11, sentenced pre 9-11	.124	.759	37.8	21	.083	.893	70.9	50	.282	.938	42.1	30
Last offense pre 9-11, sentenced post 9-11	.129	.787	40.6	24	.073	.903	75.5	54	.246	.950	41.4	27
Unconditional difference-in-difference to Whites					-.014*	-.018	1.78	1.00	-.041***	-.015	-3.44**	-6.00*
					(.008)	(.013)	(2.09)	(13.1)	(.013)	(.011)	(1.64)	(3.22)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Means, medians, differences and difference-in-differences reported. Standard errors are reported in parentheses, allowing for clustering by ethnicity-district. The outcomes shown by ethnicity are the share of Federal criminal cases that receive a downwards departure, whether any custodial sentence is awarded, the prison sentence (in months) including zero sentence lengths, and the median sentence length (in months) including zeroes. In Panel A, the full sample of 235,484 Federal cases is used (those that come up for sentencing from 10/1/1998 to 09/30/2003) and in Panel B the natural experiment sample refers to those 40,228 cases for which sentencing occurs within a 6-month window either side of 9/11/2001. In this sample, for those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The difference reported in Panel A and the difference-in-differences reported in Panel B are both relative to White defendants. The conditional differences reported in Panel A are based on a regression of sentencing outcome for individual i of ethnicity e sentenced on day t on the following covariates: the demographic characteristics of the defendant (their ethnicity, their age, highest education level, marital status and number of dependents), their type of legal counsel, the offense type where there are 31 offense types recorded, dummies for the guideline cell they are assigned to in the pre-sentence report, dummies for the Federal court district in which the case is considered, and year dummies. OLS specifications are used except for the median sentence length where a quantile regression is used. The error term is clustered by ethnicity-district.

Table 2: Ingroup Bias in the Federal CJS

Standard errors in parentheses clustered by ethnicity-district

	(1) Downward Departure	(2) Any Sentence	(3) Sentence Length	(4) Median Sentence
Black	-.002	.030***	4.27***	.916***
	(.007)	(.006)	(.789)	(.246)
Sentenced post 9-11*Black	-.013	-.010	.400	.330
	(.008)	(.008)	(.962)	(.247)
Hispanic	.022**	.057***	4.17***	.992***
	(.009)	(.007)	(.772)	(.255)
Sentenced post 9-11*Hispanic	-.038***	-.003	-.367	.632**
	(.013)	(.007)	(.739)	(.299)
Sentenced post 9-11	.006	.016**	.873**	-.026
	(.007)	(.006)	(.428)	(.152)
Controls	Yes	Yes	Yes	Yes
p-value: [post Black = post Hispanic]	.042	.262	.472	.330
Adjusted R-squared	.256	.453	0.754	-
Observations	40,228	40,228	40,228	40,228

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in Columns 1 to 6, and quantile regression estimates are shown in Columns 7 and 8. Standard errors are reported in parentheses, where these are clustered by ethnicity-district. The NE sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The dependent variable in Column 1 is a dummy for whether the case receives a downwards departure. The dependent variable in Column 2 is a dummy for whether any prison sentence is given, the dependent variable in Column 3 is the sentence length (in months) including zero sentence lengths, and the dependent variable in Column 4 is the median sentence length. In all Columns we condition on defendant ethnicity (White, Black, Hispanic), whether the case comes up post 9-11, and interactions between the two, and the following additional controls: 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 post 9-11 x Black and post 9-11 x Hispanic dummy interactions are equal against a two sided alternative.

Table 3: Offense Type and Citizenship Status

Sentencing Outcome: Downward departure
Standard errors in parentheses clustered by ethnicity-district

	(1) Drug offenses	(2) Immigration offenses	(3) All Other offenses
Black	.011	.063	-.014*
	(.012)	(.078)	(.008)
Black x Non-Citizen	.020	-.094	.031
	(.037)	(.094)	(.037)
Sentenced post 9-11*Black	.001	-.044	-.016
	(.015)	(.125)	(.010)
Sentenced post 9-11*Black x Non-Citizen	-.045	.071	-.058
	(.047)	(.148)	(.054)
Hispanic	.023	.099*	-.006
	(.016)	(.057)	(.011)
Hispanic x Non-Citizen	.063	-.182**	.031
	(.043)	(.070)	(.029)
Sentenced post 9-11*Hispanic	-.013	-.123	-.029**
	(.019)	(.093)	(.013)
Sentenced post 9-11*Hispanic x Non-Citizen	-.044	.150	-.010
	(.051)	(.112)	(.051)
Controls	Offender characteristics, defense counsel type, guideline cell dummies and Federal district dummies.		
Adjusted R-squared	.298	.350	.091
Observations	17,583	6,737	15,617

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown throughout. Standard errors are reported in parentheses, where these are clustered by ethnicity-district. The NE sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The Columns are restricted to drug, immigration and other offenses respectively. The dependent variable throughout is a dummy for whether the case receives a downwards departure. In all Columns we condition on a full set of three- and two-way interactions between defendant ethnicity (White, Black, Hispanic), defendant citizenship (where citizens are defined as being US citizens or resident/legal aliens, and non-citizens are illegal aliens, non-US citizens and those for whom alien status is unknown), and whether the case comes up post 9-11, as well as each of these three control variables alone. In all specifications the following additional controls are included: 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.

Table 4: Prosecutor Decisions at Various Stages of the Federal CJS

Standard errors in parentheses clustered by ethnicity-district

	Natural Experiment Sample	Full Sample: Initial Charges	
	(1) Substantial Assistance Departure	(2) Non-zero Statutory Minimum	(3) Statutory Minimum
Black	-.054***	.161***	20.8***
	(.010)	(.014)	(1.63)
Hispanic	-.098***	.124***	14.0***
	(.012)	(.018)	(2.04)
Sentenced post 9-11*Black	.008		
	(.013)		
Sentenced post 9-11*Hispanic	.019		
	(.011)		
Sentenced post 9-11	-.002		
	(.010)		
Controls	Offender characteristics, defense counsel type, offense type dummies, guideline cell dummies, and Federal district dummies.	As before, minus offense type and guideline cell dummies.	
p-value: [post Black = post Hispanic]	.277	-	-
p-value: [Black = Hispanic]	-	.034	.001
Adjusted R-squared	.170	.149	.147
Observations	40,228	235,484	234,114

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown. Standard errors are reported in parentheses, where these are clustered by ethnicity-district. In Columns 1 and 2, the natural experiment sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. In Columns 3 to 6, the full sample of 235,484 Federal cases is used (those that come up for sentencing from 10/1/1998 to 09/30/2003). The dependent variable in Column 1 is whether any substantial assistance departure motion is granted at sentencing. The dependent variable in Column 2 is whether the initial offense charge is such that the statutory minimum sentence is non-zero. The fully binds relative to the recommended guideline cell at sentencing. The dependent variable in Column 3 is the statutory minimum sentence. In Column 1 we condition on the following controls: 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. In Columns 2 and 3 the same controls are included except the offense type dummies and guideline cell dummies are dropped. In Column 1 the p-value at the foot of each Column is on the null that the coefficients on the post 9-11 x Black and post 9-11 x Hispanic dummy interactions are equal against a two sided alternative. In Columns 2 and 3 the p-value at the foot of each Column is on the null that the Black and Hispanic coefficients are equal against a two sided alternative.

Table 5: Pre-sentence Reports

OLS regression estimates; standard errors in parentheses clustered by ethnicity-district

	(1) Criminal History Score	(2) Offense Severity Score	(3) Minimum Guideline Sentence
Convicted and Sentenced after 9-11 [T2]*Black	.036	-.040	2.02
	(.055)	(.207)	(2.13)
Convicted and Sentenced after 9-11 [T2]*Hispanic	.016	-.625***	-2.31
	(.047)	(.221)	(1.65)
Convicted and Sentenced after 9-11 [T2]	.048	.391***	2.57**
	(.036)	(.133)	(1.28)
Convicted before 9-11 but Sentenced after 9-11 [T1]*Black	-.022	.042	.528
	(.057)	(.217)	(2.05)
Convicted before 9-11 but Sentenced after 9-11 [T1]*Hispanic	-.053	-.316	.084
	(.043)	(.218)	(1.54)
Convicted before 9-11 but Sentenced after 9-11 [T1]	-.041	-.085	-1.84
	(.034)	(.152)	(1.18)
Black	.612***	1.82***	19.8***
	(.042)	(.196)	(1.72)
Hispanic	-.337***	.545**	-1.15
	(.041)	(.225)	(1.73)
Controls	Offender characteristics, defense counsel type, offense type dummies and Federal district dummies.		
p-value: [post T1-Hispanic = post T2-Hispanic]	.225	.213	.183
Adjusted R-squared	.253	.489	.326
Observations	40,228	40,228	40,228

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in Columns 1 to 3. The natural experiment sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The dependent variable in Column 1 (2) is the criminal history score (offense severity score) reported in the pre-sentence report, and in Column 3 it is the lowest sentence in the recommended guideline cell. In all Columns we condition on defendant ethnicity (White, Black, Hispanic), whether the defendant is convicted before 9-11 but sentenced after 9-11 [treatment group T1], whether the defendant is convicted and Sentenced after 9-11 [treatment group T2], and interactions between the two treatment dummies and offender ethnicity, and the following additional controls: 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, and Federal district dummies. The p-value at the foot of each Column is on the null that the coefficients on the Convicted before 9-11 but Sentenced after 9-11 [T1]*Hispanic dummy and Convicted and Sentenced after 9-11 [T2]*Hispanic dummy interactions are equal against a two sided alternative.

Table 6: Prosecutor's Initial Offense Charges

Standard errors in parentheses clustered by ethnicity-district

	Prosecutor's Initial Charges		Judge's Sentencing		
	(1) Non-zero Statutory Minimum	(2) Statutory Minimum Length	(3) Sentence Length	(4) Sentence Length	(5) Downward Departure
Initial charges post 9-11*Hispanic	.075*	10.7**	9.33**	1.81	.017
	(.042)	(5.34)	(4.65)	(2.65)	(.024)
Hispanic	.073**	7.27	-.355	2.10	-.017
	(.034)	(5.08)	(5.04)	(2.36)	(.025)
Initial charges post 9-11*Black	-.010	.684	-5.39	.846	.019
	(.048)	(7.50)	(7.36)	(3.66)	(.025)
Black	.199***	23.3***	35.5***	5.63*	-.024
	(.033)	(4.82)	(5.69)	(2.96)	(.022)
Initial charges post 9-11	-.033	-5.96	-8.29**	-.873	-.022
	(.033)	(3.90)	(3.94)	(2.34)	(.018)
Controls	Offender characteristics, defense counsel type and Federal district dummies.			Offender characteristics, defense counsel type, offense type, guideline cell dummies and Federal district dummies.	
p-value: [post Black = post Hispanic]	.046	.172	.030	.755	.934
Adjusted R-squared	.171	.147	.190	.797	.289
Observations	3,612	3,600	3,612	3,612	3,612

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 sample of Federal cases used is: (i) for those with initial charges after 9/11, defendants in (out of) custody committed their last offense between 14 (21) days before 9/11 and the day before 9/11; (ii) for those with initial charges before 9/11, defendants in (out of) custody committed their last offense between 42 (63) days before 9/11 and 38 (42) days before 9/11. The dependent variable in Column 1 is a dummy for whether the defendant receives an initial charge with a non-zero statutory minimum sentence. The dependent variable in Column 2 is the length of statutory minimum sentence. The dependent variable in Columns 3 and 4 is the actual sentence length in months (as determined at the sentencing stage) and the dependent variable in Column 5 is a dummy for whether the case receives a downwards departure at sentencing. In Columns 1 to 3 the following controls are included: 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) and Federal district dummies. In Columns 4 and 5 the additional controls are offence type dummies and guideline cell dummies. The p-value at the foot of each Column is on the null that the coefficients on the post 9-11 x Black and post 9-11 x Hispanic dummy interactions are equal against a two sided alternative.

Table 7: Juhn-Murphy-Pierce Decompositions of Hispanic-White Differentials

	Cohort 1: Judge Decisions	Cohort 2: Prosecutor Decisions
	(1) Downwards Departure	**(2) Statutory Minimum Length**
1. Pre-9/11 (raw) differential	0.158	-1.214
2. Post-9/11 (raw) differential	0.117	6.182
3. Change in differential	-0.041	7.396
4. Due to observables: X-effect + β-effect	-0.003	-9.285
5. Due to unobservables: θ-effect + σ-effect	-0.038	16.681
6. Observable quantity: X-effect	0.005	-2.717
7. Observable penalties: β-effect	-0.008	-6.568
8. Unobservable quantities: θ-effect	-0.044	22.22
9. Unobservable penalties: σ-effect	0.006	-5.539
X-Controls	Offender characteristics, defense counsel type, offense type dummies, guideline cell dummies, and Federal district dummies.	Offender characteristics, defense counsel type and Federal district dummies.

Notes: A Juhn-Murphy-Pierce (1991) decomposition, using a non-parametric procedure, is implemented. This decomposes the unconditional difference-in-difference for each sentencing outcome between Hispanics and Whites. In Column 1 this is based on Federal criminal cases in the Natural Experiment sample. Hence the decomposition is based on 29,352 cases for Hispanic or White defendants that come up for sentencing in a six month window either side of 9/11/2001. For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The outcome in Column 1 is for whether any downward departure is received. In Column 2 the sample of Federal cases used is: (i) for those with initial charges after 9/11, defendants in (out of) custody committed their last offense between 14 (21) days before 9/11 and the day before 9/11; (ii) for those with initial charges before 9/11, defendants in (out of) custody committed their last offense between 42 (63) days before 9/11 and 38 (42) days before 9/11. The outcome in Column 2 is the length of statutory minimum sentence following from the initial offense charge. For both Juhn-Murphy-Pierce decompositions, Whites are chosen as the reference group.

Table 8: Judges and Ingroup Bias

Dependent Variable in Columns 1-4: Coefficient on post 9-11 x Hispanic x District dummy, from NE sample
Dependent Variable in Column 5: Coefficient on Hispanic x District dummy, from full sample
**Dependent Variable in Column 6: Coefficient on Black x District dummy, from full sample**
Observations weighted by districtshare of Hispanics in 2001, robust standard errors in parentheses

	Natural Experiment Sample				Full Sample	
	(1) Ethnicity	(2) Other Judge Characteristics	(3) State Population	(4) Effect Size	(5) Hispanic Coefficient, Effect Size	(6) Black Coefficient, Effect Size
District Proportion Hispanic Judges	.191***	.174*	.331**	.021**	.028*	-.009**
	(.069)	(.100)	(.140)	(.009)	(.014)	(.004)
District Proportion Black Judges	.280	.332	-.032	-.003	-.016	.001
	(.253)	(.257)	(.200)	(.017)	(.014)	(.005)
District Proportion Senior Status Judges		.030	.137*	.020*	.023	-.010
		(.079)	(.077)	(.011)	(.022)	(.007)
District Proportion Male Judges		.025	-.170*	-.020*	-.040**	.013**
		(.075)	(.091)	(.011)	(.018)	(.005)
District Proportion Democratic President Elected Judges		.196**	.153**	.028**	-.003	.001
		(.089)	(.076)	(.014)	(.013)	(.005)
State Proportion Black (2000)			.396**	.042**	-.044	.007
			(.172)	(.018)	(.031)	(.009)
State Proportion Hispanic (2000)			-.343*	-.033*	-.024	.018*
			(.198)	(.019)	(.029)	(.009)
Change in State Proportion Black (1990 - 2000)			-6.69***	-.062***	.013	-.005
			(2.25)	(.021)	(.035)	(.006)
Change in State Proportion Hispanic (1990 - 2000)			.028	.001	.013	-.017*
			(.792)	(.016)	(.029)	(.010)
Mean of Dependent Variable		-.462			-.056	-.033
Adjusted R-squared	.061	.123	.291	.291	.287	.253
Observations	90	90	90	90	90	90

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The results in Columns 1 to 4 are based on the Natural experiment sample (those that come up for sentencing in a six month window either side of 9/11/2001, where for those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The results in Columns 5 and 6 are based on the full sample (those that come up for sentencing from 10/1/1998 to 09/30/2003). Each observation represents a single Federal court district and observations are weighted by the share of Hispanics in the district in the relevant sample of Federal criminal cases (the natural experiment or full sample). Robust standard errors are reported. For Columns 1-5, the dependent variable is the coefficient on post 9-11*Hispanic*District from a difference-in-difference-in-difference regression for the Natural experiment sample period where in this first stage the full set of controls is included, and the dependent variable is whether a downwards departure is granted. In Column 5, the dependent variable is the coefficient on Hispanic*District from a difference-in-difference regression for the full sample period with a full set of controls, and where the dependent variable is whether a downward departure is granted. In Column 6, the dependent variable is the coefficient on Black*District from a difference-in-difference regression for the full sample period with a full set of controls, and where the dependent variable is whether a downward departure is granted. The data for judicial characteristics are sourced from the *Biographical Directory of Federal Judges*. In order to select the relevant judges to construct characteristics for, we used the data on commission and termination dates for each judge in the database, and in Columns 1-4 we restricted the sample to judges commissioned before the end of the natural experiment sample and those who terminated the bench after the beginning of the sample. We perform an analogous sample cut of judges relevant for the full sample in Columns 5 and 6. Data for state level characteristics are from the 1990 and 2000 5% US census data. State proportions were constructed using the individual weights (perwt) provided by IPUMS. In Columns 4 to 6, effect sizes on all covariates are reported.

Figure 1: Federal CJS Timeline

Stage: | 0. Arrest/offense | 1. Initial Appearance | 2. Bail | 3. Arraignment | 4. Initial District Court Appearance | 5. Pre-trial Motions | 6. Plea | 7. Trial | 8. Sentencing | 9. Appeals

Duration: <1 day ... <1 day ... 3-7 days ... <1 day

Days Between Stages: Maximum 14 21 days if in [out of] custody, from initial appearance ... Differs by circuit court ... 75 [90] days if in [out of] custody ... Almost no delay to state intention to appeal

Linkage Rates

Administrative Data Links: | 0. Arrest/offense | Stages 1-3: Initial Appearance through to Arraignment | Stages 4-7: Initial District Court Appearance through to Trial | Stage 8: Sentencing

Panel A. Right-to-Left Linkage Rates

Ethnicity	Offense Type			
All	All	75.1%	84.7%	90.2%
White, Black , Hispanic	All	71.8%, 70.2%, 80.8%	86%, 87.1%, 82.2%	91.4%, 91.6%, 88.4%
White, Black , Hispanic	Drug	73.8%, 68.7%, 78.3%	88.2%, 89.2%, 81.2%	92.3%, 91.9%, 88.9%
White, Black , Hispanic	Immigration	78.7%, 71.1%, 84.9%	83.4%, 79.3%, 83.5%	85.6%, 90.5%, 88.4%

Panel B. Left-to-Right Linkage Rates

Race	Offense Type			
All	All	38.2%	95.6%	84.3%
White, Black	All	37.8%, 39.3%	95.6%, 95.6%	83.7%, 86.0%
White, Black	Drug	55.1%, 53.8%	86.2%, 87.7%	86.2%, 87.7%
White, Black	Immigration	34.1%, 44.5%	81.7%, 76.2%	81.7%, 76.2%

Notes: We use the Monitoring of Federal Criminal Sentences (MFCS) data set for our analysis. This comprises information gathered from four linked administrative data sources. As described in Rehavi and Starr [2014], the four linked data sets are: (i) US Marshals Service (USMS) data, that covers the arrest/offense stage (Stage 0) and includes all persons arrested by Federal law enforcement agencies, persons arrested by local officials and then transferred to Federal custody, and persons who avoid arrest by self-surrendering; (ii) Executive Office for US Attorneys (EOUSA) data, covering initial appearance through to arraignment (Stages 1-3): these data come from the internal case database used by Federal prosecutors, and covers every case in which any prosecutor at a US Attorney's office opens a file; (iii) Administrative Office of the US Courts (AOUSC) data, covering initial district court appearances through to trial (Stages 4-7): these originate from Federal Courts and contain data on all criminal cases heard by Federal district judges, and any non-petty charge handled by a Federal magistrate judge; (iv) US Sentencing Commission (USSC) data, covering the sentencing Stage 8: this data set collects information on any case that results in conviction and sentencing for a non-petty offense. These data are collected by the Bureau of Justice Statistics.

Figure 2: Sentencing and Last Offense Dates, by Ethnicity

A. Sentencing Date   B. Date of Last Offense

Whites
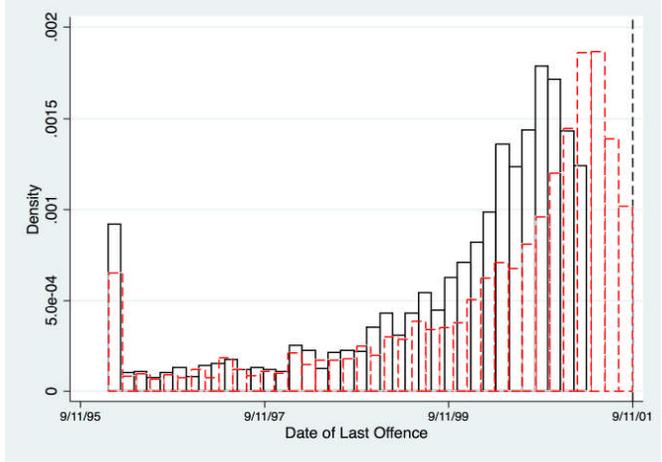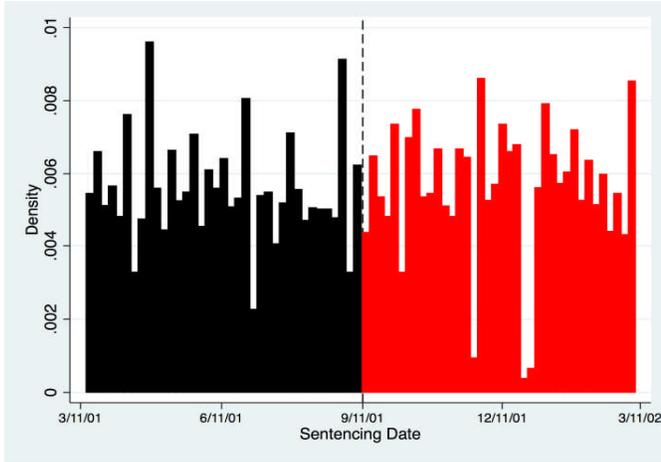
Blacks
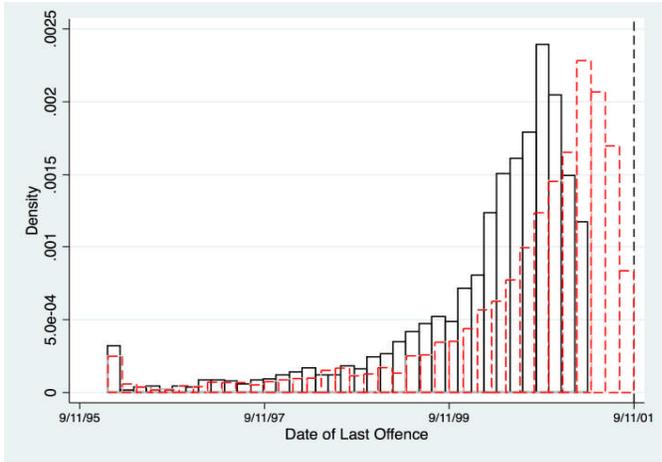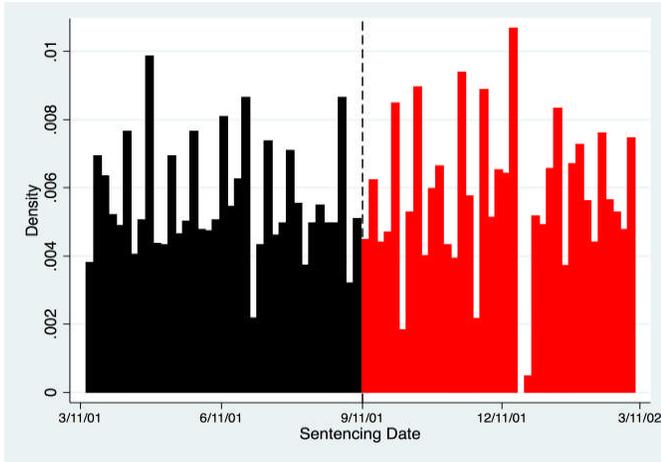
Hispanics
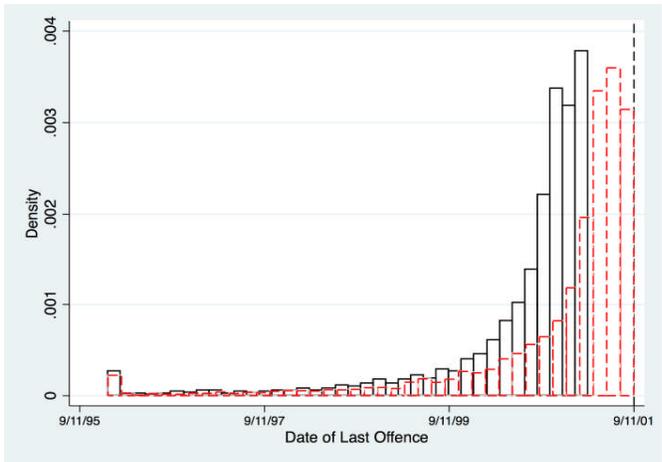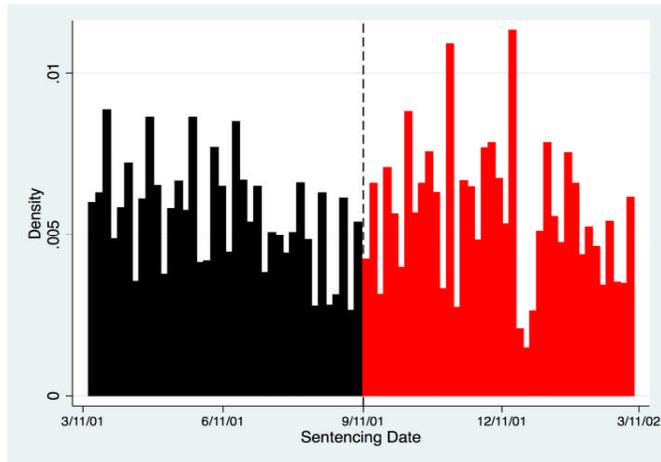
Notes: The left hand side figures show the distribution of dates of sentencing date, for each ethnicity: 9/11 is indicated by the vertical dashed line. The right hand side figures show the distribution of the dates of last offenses, by ethnicity. The first bar corresponds to a last offense date on or before 1st January 1996. The overlaid histograms are for those sentenced pre- and post-9/11. For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001.

Figure 3: Spatial Patterns of Hispanic-White Sentencing Differentials

Panel A. Share of State Population that is Hispanic, 2000 Census Data
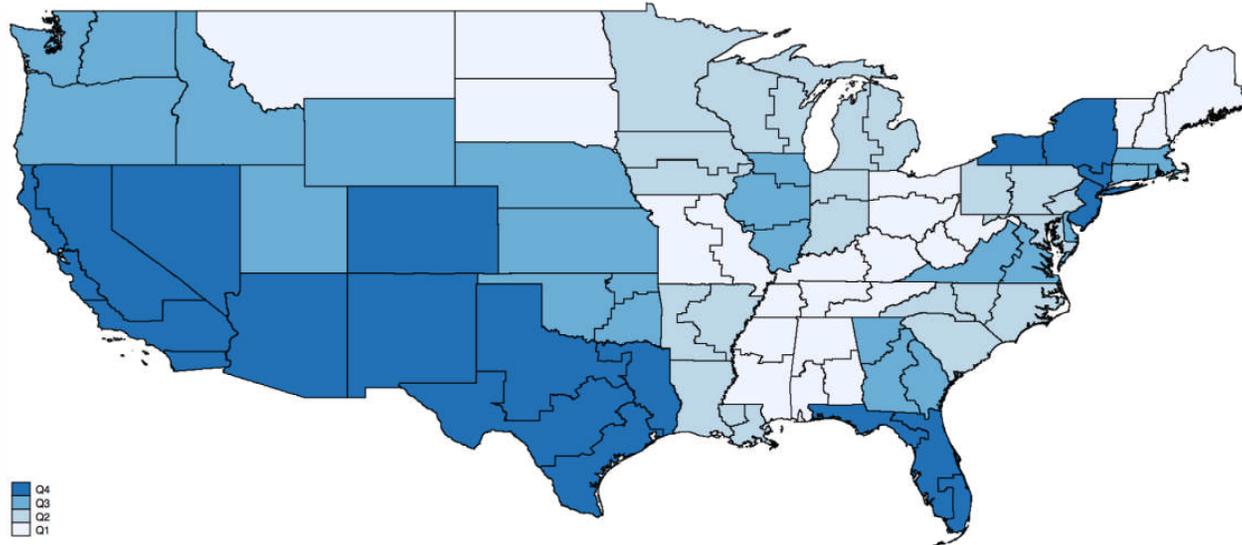
Panel B. Estimated Hispanic-White Sentencing Differential for Downward Departures Post 9-11
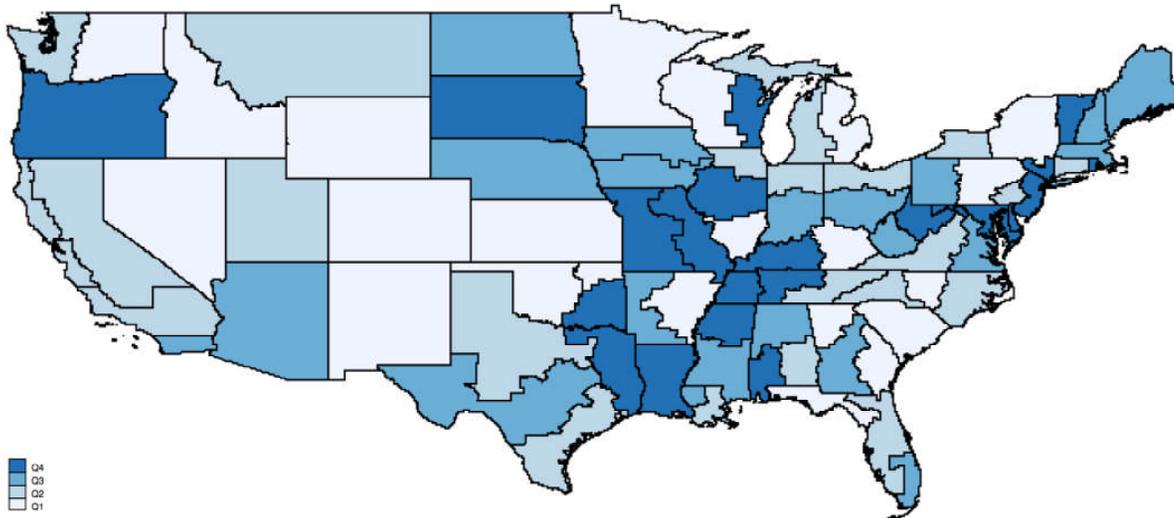
Notes: In Panel A, we use the 2000 5% US census sample to show the share of each State's population this is Hispanic, split into quartiles. In Panel B, we plot the coefficient on post 9-11*Hispanic*District from a difference-in-difference-in-difference regression for the Natural Experiment sample period where in this first stage the full set of controls is included, and the dependent variable is whether a downwards departure is granted. These coefficients are split into quartiles.

Table A1: Sentencing Guideline Cells (in months imprisonment)

	offense Level	I (0 or 1)	II (2 or 3)	III (4, 5, 6)	IV (7, 8, 9)	V (10, 11, 12)	VI (13 or more)
				Criminal History Category (Criminal History Points)			
Zone A	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
	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
Zone D	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
	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

Table A2: Detailed Federal CJS Timeline

	Stage	Who is involved	Description	Notes
1	Initial Appearance	Defendant, Federal Magistrate, Prosecutor (Assistant US Attorney), Assistant Federal Public Defender	If defendant cannot afford counsel, they fill out a financial affidavit, and are assigned to either a federal public defender or CJA panel counsel	A federal magistrate presides over proceedings until the defendant appears in district court (at Stage 4)
2	Bail	Defendant, Federal Magistrate, Prosecutor (Assistant US Attorney), defense Counsel, Pretrial Services	The bail hearing generally takes place within a week of the initial appearance, and depends on the case. Defendants seeking bail are then referred to Pretrial Services (neutral court employees, who interview the defendant and prepare a short life background and criminal history for the court). defense is present for this. Bail is then decided upon.	For "presumption" cases (drug dealing, bank robbery, child sex offenses), the govt. automatically gets 3 days to prepare for a bail hearing. If the govt. can prove the defendant is a flight risk, they get 3 days preparation time. The defense can ask for up to 5 days preparation time.
3	Arraignment	Defendant, Federal Magistrate, Prosecutor (Assistant US Attorney), defense Counsel, Federal Grand Jury	Happens within 14 (21) days from initial appearance for in-custody (out-of-custody) defendants. Defendant is arraigned on an indictment, which contains federal charges against him/her. Reviewed by grand jury. If sufficient evidence, jury "returns the indictment". After arraignment, magistrate adds the case to the district court calendar, and a district court judge is assigned. This judge will preside over the rest of the stages up to and including sentencing.	This is the stage where initial charges are filed, and so determines the statutory maximum and minimum for the offense.
4	Initial District Court Appearance	Defendant, District Court Judge, Prosecutor (Assistant US Attorney), defense Counsel	"Status" is decided: defense reviews the evidence ("discovery") in order to identify any motions. defense also discusses any pretrial dispositions (deals) with the prosecutor.	
5	Pretrial Motions	Defendant, Prosecutor (Assistant US Attorney), defense Counsel	Further prosecutor-defense interaction. The defendant's motion is sometimes called the moving papers or the opening brief. The prosecutor usually has one to three weeks to respond to the motion (the response is called an "Opposition"). The defense then typically has one or two weeks to respond to the Opposition (the defense response is called a "Reply"). One to two weeks after the Reply is filed, the court usually hears argument on the motion.	Modal pretrial motion is a suppression motion, where defense moves to suppress evidence or prevent the govt using it at trial.
6	Plea	Defendant, Prosecutor (Assistant US Attorney), defense Counsel	Guilt Plea is choice for large majority of case; either an open plea (no plea agreement) or with a plea agreement made with the prosecutor. Defense must inform defendant of every plea offer the prosecutor makes, and generally advises defendant on pros/cons of agreement. Defendant alone decides.	
7	Trial	Defendant, District Court Judge, Prosecutor (Assistant US Attorney), defense Counsel, Jury	The typical federal trial lasts 3-7 days. At the trial, the defendant has the right to testify – or to not testify, and if he or she does not testify, that cannot be held against the defendant by the jury. The defendant also has the right to "confront" (i.e., cross-examine) government witnesses, and can use the subpoena power of the court to secure evidence or witnesses for trial.	
8	Sentencing	Defendant, District Court Judge, Prosecutor (Assistant US Attorney), defense Counsel, Probation Office	If a defendant is convicted, sentencing takes place 75 (90) days later if the defendant is in (out of) custody. A defendant convicted of some offenses will likely be remanded into custody after trial. After a conviction, the defendant and his or her attorney complete forms relating to the defendant's life history and provide those to the (neutral) Probation Office. Several weeks after the conviction, the defendant will be interviewed by a Probation Officer, with defense counsel present. The Probation Officer will then take information from that interview, from the forms submitted by the defense, and from material provided by the government, and will prepare a draft presentence report. The draft presentence report (or PSR) is provided to defense counsel and the government 35 days before sentencing. The parties must make factual or legal objections to the report within 10 days of receipt. 14 days before sentencing, the final PSR is provided to the judge. This final PSR describes the defendant's background, describes the offense, and calculates the federal sentencing guidelines. It also includes a recommended sentence, and lists any unresolved objections. 7 days before sentencing, the parties submit sentencing memoranda to the court, arguing for their proposed sentences. 3 days later, the parties may submit replies to the sentencing memos. At the sentencing hearing, the district court judge must resolve any remaining objections to the PSR, make factual findings, and must consider the factors of the key sentencing statute, 18 USC § 3553(a). Before imposing the sentence, the court must permit the defendant to speak (or "allocute").	
9	Appeals	Defendant, District Court Judge, Supreme Court Judge	If the defendant did not waive the right to appeal in a plea agreement, the defense may appeal both the conviction and the sentence imposed. The public defender will continue to represent the defendant, for free, during the appeal. If the defendant does not win the appeal in their Circuit, he or she can file a petition for writ of certiorari with the Supreme Court of the United States. The public defender will continue to represent the defendant during the petition for certiorari and Supreme Court argument, if the writ is granted.	There is a very short period during which the defense must state its intention to appeal ("notice" an appeal), so the subject should be discussed immediately after sentencing.

Source: http://gan.fd.org/pdfs/NDGA%20Timeline.pdf, accessed March 7th 2016.

Table A3a: Descriptives for the Full Sample

Means, standard deviations in parentheses, p-values in brackets

	White			Black			Hispanic			Total		
	Raw Sample	Working Sample	p-value	Raw Sample	Working Sample	p-value	Raw Sample	Working Sample	p-value	Raw Sample	Working Sample	p-value
Sample Size	75931	73786		62384	60653		111973	101045		250288	235484	
Number Dependents	1.146	1.147	[0.991]	1.671	1.675	[0.939]	1.844	1.846	[0.984]	1.577	1.576	[0.990]
	(1.435)	(1.434)		(1.841)	(1.842)		(1.8)	(1.796)		(1.731)	(1.729)	
Marital Status:												
Single	0.336	0.337	[0.924]	0.535	0.536	[0.937]	0.311	0.328	[0.520]	0.374	0.384	[0.641]
	(0.472)	(0.473)		(0.499)	(0.499)		(0.463)	(0.469)		(0.484)	(0.486)	
Married	0.351	0.353	[0.89]	0.205	0.205	[0.906]	0.324	0.344	[0.380]	0.303	0.311	[0.520]
	(0.477)	(0.478)		(0.403)	(0.404)		(0.468)	(0.475)		(0.459)	(0.463)	
Cohabiting	0.072	0.073	[0.93]	0.121	0.122	[0.923]	0.139	0.146	[0.617]	0.114	0.117	[0.783]
	(0.259)	(0.26)		(0.326)	(0.327)		(0.345)	(0.353)		(0.318)	(0.321)	
Divorced	0.161	0.162	[0.893]	0.063	0.064	[0.904]	0.048	0.05	[0.678]	0.086	0.088	[0.776]
	(0.367)	(0.368)		(0.243)	(0.244)		(0.213)	(0.218)		(0.28)	(0.284)	
Widowed	0.006	0.006	[0.850]	0.003	0.003	[0.970]	0.002	0.002	[0.626]	0.003	0.004	[0.716]
	(0.076)	(0.076)		(0.053)	(0.053)		(0.048)	(0.049)		(0.059)	(0.06)	
Separated	0.049	0.049	[0.900]	0.049	0.049	[0.924]	0.044	0.046	[0.628]	0.046	0.048	[0.610]
	(0.215)	(0.216)		(0.216)	(0.216)		(0.204)	(0.209)		(0.21)	(0.213)	
Education Level:												
Less than High School	0.258	0.26	[0.915]	0.403	0.404	[0.952]	0.596	0.635	[0.384]	0.445	0.458	[0.744]
	(0.438)	(0.438)		(0.491)	(0.491)		(0.491)	(0.481)		(0.497)	(0.498)	
High School Graduate	0.375	0.377	[0.876]	0.364	0.365	[0.901]	0.158	0.165	[0.739]	0.275	0.283	[0.779]
	(0.484)	(0.485)		(0.481)	(0.481)		(0.365)	(0.371)		(0.447)	(0.45)	
Some College	0.225	0.226	[0.905]	0.183	0.184	[0.887]	0.07	0.072	[0.925]	0.145	0.149	[0.839]
	(0.418)	(0.418)		(0.386)	(0.387)		(0.256)	(0.258)		(0.352)	(0.356)	
College Graduate	0.125	0.125	[0.961]	0.037	0.038	[0.971]	0.019	0.019	[0.982]	0.056	0.057	[0.886]
	(0.331)	(0.331)		(0.19)	(0.19)		(0.137)	(0.137)		(0.23)	(0.232)	
Age	38.517	38.556	[0.927]	31.886	31.912	[0.926]	32.201	32.146	[0.907]	34.076	34.11	[0.949]
	(12.202)	(12.191)		(9.299)	(9.291)		(9.200)	(9.220)		(10.671)	(10.705)	
Defense Counsel:												
Privately Retained	0.167	0.169	[0.974]	0.079	0.08	[0.961]	0.072	0.072	[0.98]	0.102	0.104	[0.919]
	(0.373)	(0.374)		(0.269)	(0.271)		(0.258)	(0.259)		(0.303)	(0.306)	
Court Appointed	0.172	0.173	[0.981]	0.174	0.176	[0.974]	0.297	0.298	[0.998]	0.228	0.227	[0.978]
	(0.377)	(0.378)		(0.379)	(0.38)		(0.457)	(0.457)		(0.42)	(0.419)	
Federal Public Defender	0.122	0.122	[0.979]	0.142	0.141	[0.971]	0.26	0.262	[0.97]	0.188	0.187	[0.963]
	(0.327)	(0.328)		(0.349)	(0.348)		(0.439)	(0.44)		(0.391)	(0.39)	
Self-represented	0.004	0.004	[0.798]	0.003	0.002	[0.917]	0.001	0.000	[0.58]	0.002	0.002	[0.811]
	(0.066)	(0.063)		(0.05)	(0.049)		(0.024)	(0.021)		(0.047)	(0.045)	
Rights waived	0.003	0.003	[0.779]	0.003	0.003	[0.836]	0.001	0.001	[0.761]	0.002	0.002	[0.752]
	(0.056)	(0.053)		(0.057)	(0.054)		(0.034)	(0.032)		(0.048)	(0.046)	
Other Arrangements	0.000	0.000	[0.998]	0.001	0.001	[0.979]	0.000	0.000	[0.924]	0.000	0.001	[0.925]
	(0.022)	(0.022)		(0.024)	(0.024)		(0.02)	(0.021)		(0.022)	(0.022)	
Criminal History Score	2.181	2.186	[0.939]	3.041	3.054	[0.857]	2.42	2.404	[0.882]	2.503	2.503	[0.997]
	(1.639)	(1.641)		(1.836)	(1.836)		(1.683)	(1.666)		(1.742)	(1.737)	
Offense Severity	17.859	17.896	[0.903]	22.166	22.216	[0.900]	18.444	18.399	[0.958]	19.21	19.225	[0.977]
	(8.470)	(8.454)		(9.452)	(9.43)		(8.125)	(8.050)		(8.762)	(8.732)	

Notes: The full sample refers to all Federal cases that come up for sentencing from 10/1/1998 to 09/30/2003. For each ethnicity (and the sample as a whole), we show the descriptive statistic for all these cases (the "Raw Sample" Columns), and for those cases used in the main analysis where there is non-missing information for key covariates (the "Working Sample" Columns). Specifically, observations were dropped from the raw sample if the following variables were missing: district, race/ethnicity, criminal history, offense severity, sentence length or offense type. Means and standard deviations (in parentheses) are shown. The p-values are tests of equality of the statistic within ethnic group across the two samples, based on an OLS regression that allows standard errors to be clustered by ethnicity-district.

Table A3b: Sample Descriptives

Means, standard deviations in parentheses, p-values in brackets

	White			Black			Hispanic			Total		
	Raw Sample	Working Sample	p-value	Raw Sample	Working Sample	p-value	Raw Sample	Working Sample	p-value	Raw Sample	Working Sample	p-value
Sample Size	14226	12994		12054	10876		18212	16358		44492	40228	
Number Dependents	1.102	1.101	[0.977]	1.685	1.689	[0.938]	1.857	1.839	[0.851]	1.561	1.555	[0.934]
	(1.42)	(1.415)		(1.823)	(1.831)		(1.792)	(1.776)		(1.72)	(1.713)	
Marital Status:												
Single	0.332	0.336	[0.650]	0.53	0.534	[0.762]	0.319	0.328	[0.678]	0.380	0.387	[0.738]
	(0.471)	(0.473)		(0.499)	(0.499)		(0.466)	(0.47)		(0.485)	(0.487)	
Married	0.359	0.356	[0.828]	0.208	0.206	[0.793]	0.344	0.355	[0.547]	0.312	0.315	[0.819]
	(0.48)	(0.479)		(0.406)	(0.404)		(0.475)	(0.478)		(0.463)	(0.464)	
Cohabiting	0.078	0.077	[0.966]	0.128	0.128	[0.970]	0.153	0.155	[0.877]	0.122	0.123	[0.935]
	(0.268)	(0.267)		(0.334)	(0.334)		(0.36)	(0.362)		(0.327)	(0.328)	
Divorced	0.159	0.159	[0.975]	0.061	0.061	[0.953]	0.052	0.052	[0.959]	0.089	0.089	[0.962]
	(0.365)	(0.366)		(0.239)	(0.239)		(0.223)	(0.222)		(0.284)	(0.285)	
Widowed	0.005	0.005	[0.905]	0.003	0.003	[0.591]	0.002	0.002	[0.881]	0.003	0.003	[0.734]
	(0.068)	(0.067)		(0.054)	(0.051)		(0.046)	(0.046)		(0.056)	(0.055)	
Separated	0.048	0.048	[0.960]	0.048	0.048	[0.950]	0.047	0.047	[0.983]	0.048	0.047	[0.981]
	(0.214)	(0.214)		(0.214)	(0.213)		(0.211)	(0.211)		(0.213)	(0.213)	
Education Level:												
Less than High School	0.262	0.265	[0.810]	0.403	0.405	[0.852]	0.613	0.634	[0.529]	0.444	0.453	[0.799]
	(0.440)	(0.441)		(0.49)	(0.491)		(0.487)	(0.482)		(0.497)	(0.498)	
High School Graduate	0.383	0.384	[0.950]	0.372	0.371	[0.933]	0.182	0.183	[0.938]	0.298	0.299	[0.960]
	(0.486)	(0.486)		(0.483)	(0.483)		(0.386)	(0.387)		(0.457)	(0.458)	
Some College	0.22	0.22	[0.982]	0.181	0.181	[0.940]	0.076	0.074	[0.884]	0.15	0.15	[0.991]
	(0.414)	(0.414)		(0.385)	(0.385)		(0.264)	(0.262)		(0.357)	(0.357)	
College Graduate	0.124	0.122	[0.830]	0.038	0.037	[0.892]	0.022	0.02	[0.778]	0.059	0.057	[0.874]
	(0.33)	(0.327)		(0.191)	(0.189)		(0.146)	(0.14)		(0.235)	(0.233)	
Age	38.439	38.206	[0.541]	31.797	31.718	[0.787]	32.325	32.162	[0.723]	34.143	33.997	[0.769]
	(12.167)	(12.093)		(9.251)	(9.242)		(9.285)	(9.234)		(10.707)	(10.655)	
Defense Counsel:												
Privately Retained	0.165	0.166	[0.973]	0.079	0.08	[0.949]	0.082	0.079	[0.887]	0.108	0.107	[0.985]
	(0.371)	(0.372)		(0.27)	(0.272)		(0.275)	(0.27)		(0.31)	(0.31)	
Court Appointed	0.170	0.170	[1]	0.159	0.16	[0.963]	0.275	0.274	[0.994]	0.21	0.21	[0.998]
	(0.376)	(0.376)		(0.365)	(0.367)		(0.446)	(0.446)		(0.407)	(0.407)	
Federal Public Defender	0.132	0.134	[0.905]	0.152	0.154	[0.952]	0.256	0.267	[0.852]	0.188	0.194	[0.869]
	(0.338)	(0.341)		(0.359)	(0.361)		(0.437)	(0.442)		(0.391)	(0.395)	
Self-represented	0.004	0.003	[0.576]	0.003	0.002	[0.646]	0.000	0.000	[0.718]	0.002	0.002	[0.487]
	(0.061)	(0.054)		(0.056)	(0.047)		(0.02)	(0.017)		(0.047)	(0.041)	
Rights waived	0.001	0.001	[0.819]	0.002	0.002	[0.987]	0.001	0.001	[0.930]	0.001	0.001	[0.951]
	(0.032)	(0.034)		(0.047)	(0.047)		(0.029)	(0.029)		(0.036)	(0.036)	
Other Arrangements	0.000	0.000	[0.948]	0.000	0.000	[0.942]	-	-	-	0.000	0.000	[0.924]
	(0.012)	(0.012)		(0.018)	(0.019)		-	-		(0.012)	(0.012)	
Criminal History Score	2.209	2.214	[0.944]	3.067	3.061	[0.934]	2.377	2.39	[0.900]	2.511	2.514	[0.961]
	(1.66)	(1.657)		(1.832)	(1.822)		(1.673)	(1.667)		(1.748)	(1.741)	
Offense Severity	18.107	17.81	[0.311]	22.119	21.705	[0.304]	19.307	18.982	[0.671]	19.687	19.34	[0.395]
	(8.358)	(8.205)		(9.209)	(9.025)		(7.971)	(7.695)		(8.594)	(8.376)	

Notes: The natural experiment sample refers to all cases for which sentencing occurs within a 6-month window of 9/11/2001. For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. For each ethnicity (and the sample as a whole), we show the descriptive statistic for all these cases (the "Raw Sample" columns), and for those cases used in the main analysis where there is non-missing information for key covariates (the "Working Sample" Columns). Specifically, observations were dropped from the raw sample if the following variables were missing: district, race/ethnicity, criminal history, offense severity, sentence length, offense type or date of final offense. We further restrict the sample to cases in which: (i) guilt pleas are filed (that is so for 96% of defendants); (ii) three or fewer offenses were committed because for offenses in the 2002 tax year (those that come up for sentencing from 01/10/2001 through to 30/09/2002), in the MCFS data we only observe the date of offense for the first three offenses. Means and standard deviations (in parentheses) are shown. The p-values are tests of equality of the statistic within ethnic group across the two samples, based on an OLS regression that allows standard errors to be clustered by ethnicity-district.

Table A4: Robustness Checks

Standard errors in parentheses, clustered by ethnicity-district unless otherwise stated

	Cluster on sentence week x ethnicity				Excluding Cases Where Statutory Minima or Maxima Bind Partially			
	(1a) Downwards Departure	(1b) Any Sentence	(1c) Sentence	(1d) Median Sentence	(2a) Downwards Departure	(2b) Any Sentence	(2c) Sentence	(2d) Median Sentence
Black	-0.002	0.030***	4.272***	0.916***	-0.004	0.033***	4.083***	0.569***
	(0.007)	(0.005)	(0.671)	(0.181)	(0.007)	(0.006)	(0.734)	(0.208)
Sentenced post 9-11*Black	-0.013	-0.010	0.400	0.330	-0.016*	-0.009	-0.833	0.139
	(0.008)	(0.007)	(0.924)	(0.220)	(0.008)	(0.009)	(0.938)	(0.178)
Hispanic	0.022**	0.057***	4.174***	0.992***	0.020**	0.063***	3.026***	0.699***
	(0.009)	(0.006)	(0.629)	(0.174)	(0.009)	(0.007)	(0.678)	(0.223)
Sentenced post 9-11*Hispanic	-0.038***	-0.003	-0.367	0.632***	-0.041***	-0.005	0.446	0.581*
	(0.011)	(0.006)	(0.657)	(0.174)	(0.011)	(0.008)	(0.672)	(0.340)
Sentenced post 9-11	0.006	0.016***	0.873*	-0.026	0.009	0.018***	0.434	-0.036
	(0.006)	(0.005)	(0.454)	(0.124)	(0.007)	(0.007)	(0.437)	(0.111)
p-value: [post Black = post Hispanic]	0.022	0.224	0.413	0.166	0.018	0.519	0.190	0.160
Controls	Offender characteristics, defense counsel type, offense type dummies, guideline cell dummies, and Federal district dummies.							
Adjusted R-squared	0.256	0.453	0.754	-	0.275	0.464	0.809	-
Observations	40,228	40,228	40,228	40,228	32,430	32,430	32,430	32,430

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in Columns 1a to 1c, 2a to 2c, quantile regression estimates are shown in Columns 1d and 2d. Standard errors are reported in parentheses, where these are clustered by sentence week x ethnicity in Columns 1a-1d, and by ethnicity-district otherwise. The natural experiment sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The dependent variable in Columns 1a and 2a is a dummy for whether the case receives a downwards departure. The dependent variable in Columns 1b and 2b is a dummy for whether any prison sentence is given, the dependent variable in Columns 1c and 2c are the sentence length (in months) including sentences of zero length, and the dependent variable in Columns 1d and 2d is the median sentence length. In Columns 2a-2d we exclude cases where statutory minima or maxima bind partially, namely if a statutory minimum is above the lower limit of the guideline cell or when the statutory maximum is below the upper limit. In all Columns we condition on defendant ethnicity (White, Black, Hispanic), whether the defendant is sentenced after 9-11 and interactions between this treatment dummies and offender ethnicity, and the following controls: 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, and Federal district dummies. The p-value at the foot of each Column is on the null that the coefficients on the post 9-11 x Black and post 9-11 x Hispanic dummy interactions are equal against a two sided alternative.

Table A5: Sentencing Impacts of the Natural Experiment, by Ethnicity

Standard errors in parentheses clustered by ethnicity-district

	White				Black				Hispanic			
	(1a) Downwards Departure	(1b) Any Sentence	(1c) Sentence	(1d) Median Sentence	(2a) Downwards Departure	(2b) Any Sentence	(2c) Sentence	(2d) Median Sentence	(3a) Downwards Departure	(3b) Any Sentence	(3c) Sentence	(3d) Median Sentence
Sentenced post 9-11	0.004	0.015**	0.800*	0.025	-0.008	0.005	0.586	0.217	-0.030***	0.016***	0.449	0.473**
	(0.006)	(0.006)	(0.456)	(0.154)	(0.005)	(0.005)	(0.833)	(0.207)	(0.011)	(0.004)	(0.513)	(0.225)
Difference with Whites					-0.011	-0.010	-0.214	0.191	-0.034***	0.000	-0.351	0.448
					(0.008)	(0.008)	(0.950)	(0.259)	(0.013)	(0.007)	(0.686)	(0.273)
Difference with Blacks									-0.023*	0.011*	-0.137	0.257
									(0.012)	(0.006)	(0.978)	(0.306)
Controls	Offender characteristics, defense counsel type, offense type dummies, guideline cell dummies, and Federal district dummies.											
Adjusted R-squared	0.151	0.473	0.764	-	0.074	0.513	0.744	-	0.313	0.296	0.759	-
Observations	12,994	12,994	12,994	12,994	10,876	10,876	10,876	10,876	16,358	16,358	16,358	16,358

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in Columns 1a to 1c, 2a to 2c, quantile regression estimates are shown in Columns 1d and 2d. Standard errors are reported in parentheses, where these are clustered by sentence week x ethnicity in Columns 1a-1d, and by ethnicity-district otherwise. The natural experiment sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. In Columns 1a-1d only criminal cases involving White defendants are used. In Columns 2a-2d only criminal cases involving Black defendants are used. In Columns 3a-3d only criminal cases involving Hispanic defendants are used. The dependent variable in Columns 1a, 2a and 3a is a dummy for whether the case receives a downwards departure. The dependent variable in Columns 1b, 2b and 3b is a dummy for whether any prison sentence is given, the dependent variable in Columns 1c, 2c and 3c are the sentence length (in months) including sentences of zero length, and the dependent variable in Columns 1d, 2d and 3d is the median sentence length. In all Columns we condition on whether the defendant is sentenced after 9-11 and the following controls: 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, and Federal district dummies. In Columns 2a-2c we report differences between the coefficient estimate between Blacks and Whites (and the corresponding standard error). In Columns 3a-3c we report differences between the coefficient estimate between Hispanics and Whites, and Hispanics and Blacks (and the corresponding standard error).

Table A6: Racial Sentencing Differentials in the Federal CJS

Standard errors in parentheses clustered by ethnicity-district

	(1) Downwards Departure	(2) Any Sentence	(3) Sentence	(4) Median Sentence
Black	-0.017***	0.009*	2.939***	0.658**
	(0.006)	(0.005)	(0.753)	(0.263)
Sentenced post 9-11*Black	0.009	-0.008	0.196	-0.154
	(0.009)	(0.006)	(0.894)	(0.227)
American Indian	-0.042	0.015	0.496	0.526
	(0.036)	(0.020)	(2.443)	(0.704)
Sentenced post 9-11*American Indian	-0.037	-0.004	4.829*	0.615
	(0.025)	(0.018)	(2.644)	(0.729)
Asian/Pacific Islander	-0.036*	0.032	0.766	0.924**
	(0.018)	(0.024)	(0.919)	(0.443)
Sentenced post 9-11*Asian/Pacific Island	0.034	-0.022	-2.503	-1.011**
	(0.023)	(0.026)	(1.612)	(0.439)
Multi-Racial	0.074	0.026	-4.229	-5.992*
	(0.076)	(0.048)	(4.636)	(3.473)
Sentenced post 9-11*Multi-Racial	0.004	-0.023	7.285	5.621
	(0.096)	(0.079)	(7.364)	(3.817)
Other Race	0.110	-0.009	9.402**	1.480*
	(0.134)	(0.086)	(4.096)	(0.898)
Sentenced post 9-11*Other Race	-0.118	0.139	-6.955	0.852
	(0.145)	(0.144)	(5.718)	(2.551)
Sentenced post 9-11	-0.016**	0.014***	0.747*	0.381**
	(0.008)	(0.003)	(0.392)	(0.174)
Controls	Offender characteristics, defense counsel type, offense type dummies, guideline cell dummies, and Federal district dummies.			
Adjusted R-squared	0.254	0.456	0.756	-
Unadjusted R-squared	-	-	-	0.734
Observations	40,858	40,858	40,858	40,858

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in Columns 1 to 3, and quantile regression estimates are shown in Column 4. Standard errors are reported in parentheses, where these are clustered by ethnicity-district. The natural experiment sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The dependent variable in Column 1 is a dummy for whether the case receives a downwards departure. The dependent variable in Column 2 is a dummy for whether any prison sentence is given, the dependent variable in Column 3 is the sentence length (in months) including sentences of zero length, and the dependent variable in Column 4 is the median sentence length. In all Columns we condition on defendant race, whether the case comes up post 9-11, and interactions between the two, and all the following additional controls are included: 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.

Table A7: Time in the Federal CJS

Standard errors in parentheses clustered by ethnicity-district

	Include Dummies for 20 Groupings of Time Between Last Offense and Sentence Date				Include Dummies for 20 Groupings of Last Offense Date			
	(1a) Downwards Departure	(1b) Any Sentence	(1c) Sentence	(1d) Median Sentence	(2a) Downwards Departure	(2b) Any Sentence	(2c) Sentence	(2d) Median Sentence
Black	-0.002	0.029***	4.075***	1.069***	-0.001	0.029***	4.151***	0.966***
	(0.006)	(0.006)	(0.765)	(0.242)	(0.007)	(0.006)	(0.759)	(0.237)
Sentenced post 9-11*Black	-0.013	-0.009	0.426	0.217	-0.014*	-0.011	0.269	0.255
	(0.008)	(0.008)	(0.955)	(0.266)	(0.008)	(0.008)	(0.951)	(0.253)
Hispanic	0.018**	0.051***	3.734***	1.043***	0.023**	0.055***	3.800***	0.995***
	(0.009)	(0.006)	(0.764)	(0.241)	(0.009)	(0.006)	(0.757)	(0.227)
Sentenced post 9-11*Hispanic	-0.035***	-0.001	-0.367	0.514	-0.042***	-0.007	-0.566	0.490*
	(0.013)	(0.007)	(0.745)	(0.358)	(0.012)	(0.007)	(0.782)	(0.273)
Sentenced post 9-11	0.006	0.016***	0.952**	0.137	-0.002	-0.001	-0.588	-0.317*
	(0.007)	(0.006)	(0.428)	(0.169)	(0.007)	(0.006)	(0.464)	(0.172)
p-value: [post black = post Hispanic]	0.085	0.200	0.455	0.407	0.015	0.508	0.456	0.418
Controls	Offender characteristics, defense counsel type, offense type dummies, guideline cell dummies, and Federal district dummies.							
Adjusted R-squared	0.261	0.459	0.756	-	0.257	0.458	0.756	-
Observations	40,228	40,228	40,228	40,228	40,228	40,228	40,228	40,228

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in Columns 1a to 1c, 2a to 2c, quantile regression estimates are shown in Columns 1d and 2d. Standard errors are reported in parentheses, where these are clustered by sentence week x ethnicity in Columns 1a-1d, and by ethnicity-district otherwise. The natural experiment sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The dependent variable in Columns 1a and 2a is a dummy for whether the case receives a downwards departure. The dependent variable in Columns 1b and 2b is a dummy for whether any prison sentence is given, the dependent variable in Columns 1c and 2c are the sentence length (in months) including sentences of zero length, and the dependent variable in Columns 1d and 2d is the median sentence length. In all Columns we condition on defendant ethnicity (White, Black, Hispanic), whether the defendant is sentenced after 9-11 and interactions between this treatment dummies and offender ethnicity, and the following controls: 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, and Federal district dummies. In Columns 1a-1d we additionally include dummies to group the days between last offense and sentencing date into 20 bins, and in Columns 2a-2d we instead additionally include dummies to group the date of last offense into 20 bins. The p-value at the foot of each Column is on the null that the coefficients on the post 9-11 x Black and post 9-11 x Hispanic dummy interactions are equal against a two sided alternative.

Table A8: Time Between Dates of Last offense and Sentencing

OLS and survival regression estimates; standard errors in parentheses, clustered by ethnicity-district

	All Offenses			Drug Offenses			Immigration Offenses			Other Offenses		
	(1a) OLS	(1b) Cox	(1c) Log logistic, Gamma Frailty	(2a) OLS	(2b) Cox	(2c) Log logistic, Gamma Frailty	(3a) OLS	(3b) Cox	(3c) Log logistic, Gamma Frailty	(4a) OLS	(4b) Cox	(4c) Log logistic, Gamma Frailty
Sentenced post 9-11	5.955	-0.024	0.007	3.443	-0.037	0.010	-61.443	0.072	-0.006	10.796	-0.047*	0.018
	(11.228)	(0.020)	(0.016)	(15.687)	(0.045)	(0.021)	(37.545)	(0.090)	(0.055)	(14.081)	(0.025)	(0.018)
Sentenced post 9-11*Black	13.895	-0.021	0.022	12.215	0.003	0.001	84.703	-0.033	0.047	16.355	-0.034	0.037
	(14.512)	(0.029)	(0.020)	(20.557)	(0.053)	(0.029)	(66.065)	(0.202)	(0.099)	(19.476)	(0.039)	(0.025)
Sentenced post 9-11*Hispanic	8.064	-0.036	0.033	12.705	-0.074	0.035	64.367*	-0.078	0.035	19.318	0.014	0.023
	(12.431)	(0.030)	(0.022)	(17.734)	(0.056)	(0.026)	(38.723)	(0.097)	(0.058)	(26.404)	(0.062)	(0.035)
Black	-41.053***	0.038	-0.045**	-2.150	-0.021	0.027	-68.535	0.084	-0.128*	-61.788***	0.109***	-0.082***
	(10.615)	(0.029)	(0.019)	(12.463)	(0.042)	(0.023)	(47.816)	(0.175)	(0.075)	(14.977)	(0.034)	(0.021)
Hispanic	-55.377***	0.178***	-0.148***	-38.960***	0.230***	-0.110***	-61.875*	0.166**	-0.076**	-85.268***	0.145***	-0.195***
	(11.836)	(0.039)	(0.021)	(12.308)	(0.043)	(0.021)	(31.980)	(0.084)	(0.038)	(22.600)	(0.045)	(0.033)
p-value: [post Black = post Hispanic]	0.590	0.617	0.574	0.975	0.052	0.175	0.720	0.806	0.881	0.910	0.451	0.700
Controls	Offender characteristics, defense counsel type, guideline cell dummies, and Federal district dummies. offense type dummies are only controlled for in Columns 1a-1c.											
Observations	40,228	40,228	40,228	17,722	17,722	17,722	6,790	6,790	6,790	15,716	15,716	15,716

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample of cases refers to those 40,228 cases for which sentencing occurs within a 6-month window of 9/11/2001. For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. In Columns 1a-1c, the full natural experiment sample is used. In Columns 2a-2c (3a-3c) (4a-4c) the sample is restricted to drug (immigration) (other) offenses. The dependent variable is the number of days between the date of the last offense and the sentencing date. In Columns 1a, 2a, 3a and 4a an OLS model is estimated. In Columns 1b, 2b, 3b and 4b a Cox proportional hazard model is estimated so that a negative coefficient means a lower hazard rate, and thus a longer duration. In Columns 1c, 2c, 3c and 4c a log-logistic model with a frailty parameter is estimated. In this model a positive coefficient implies a longer duration. In all Columns we condition on defendant ethnicity (White, Black, Hispanic), whether the defendant is sentenced after 9-11 and interactions between this treatment dummies and offender ethnicity, and the following controls: 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); and Federal district dummies. offense type dummies are only controlled for in Columns 1a-1c. The p-value at the foot of each Column is on the null that the coefficients on the post 9-11 x Black and post 9-11 x Hispanic dummy interactions are equal against a two sided alternative.

Table A9: Placebos

Standard errors in parentheses clustered by ethnicity-district

	(1) Downwards Departure	(2) Any Sentence	(3) Sentence	(4) Median Sentence
Sentenced post 9-11	-0.003	-0.003	0.015	-0.066
	(0.004)	(0.005)	(0.417)	(0.107)
Sentenced post 9-11*Black	0.002	0.003	0.304	0.217
	(0.005)	(0.007)	(0.694)	(0.174)
Sentenced post 9-11*Hispanic	0.008	-0.001	-0.018	0.033
	(0.006)	(0.006)	(0.497)	(0.139)
Sentenced post 9-11*2001	0.008	0.020***	1.082*	0.127
	(0.008)	(0.007)	(0.621)	(0.191)
Sentenced post 9-11*Black*2001	-0.016	-0.013	0.037	0.022
	(0.010)	(0.009)	(1.097)	(0.323)
Sentenced post 9-11*Hispanic*2001	-0.047***	-0.004	-0.416	0.334
	(0.016)	(0.008)	(0.927)	(0.295)
2001	-0.001	-0.014**	-0.018	-0.013
	(0.006)	(0.006)	(0.438)	(0.128)
2001*Black	0.005	0.008	1.578*	0.213
	(0.007)	(0.007)	(0.817)	(0.251)
2001*Hispanic	0.043***	-0.004	-0.271	-0.287
	(0.015)	(0.007)	(0.635)	(0.221)
Black	-0.010	0.026***	3.346***	0.798***
	(0.007)	(0.005)	(0.699)	(0.213)
Hispanic	-0.006	0.063***	3.922***	1.390***
	(0.014)	(0.006)	(0.638)	(0.234)
Common Impact: POST*2001 - POST	.011	.023**	1.066	0.193
	(0.011)	(0.011)	(0.960)	(0.266)
Confidence Interval	[-0.010, 0.032]	[0.002, 0.043]	[-0.825, 2.957]	[-0.329, 0.714]
Controls	Offender characteristics, defense counsel type, offense type dummies, guideline cell dummies, and Federal district dummies.			
Adjusted R-squared	0.243	0.461	0.753	-
Unadjusted R-squared	-	-	-	0.732
Observations	114,642	114,642	114,642	114,642

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in Columns 1 to 3, and quantile regression estimates are shown in Column 4. Standard errors are reported in parentheses, where these are clustered by ethnicity-district. The sample of cases used are those 114,642 cases for which sentencing occurs within a 6-month window of 9/11 in years 1998 to 2001. For those defendants sentenced after 9/11 each year, the last offense was committed prior to 9/11 that year, and if sentenced before 9/11 each year, the last offense was committed at least 180 days prior to 9/11 that year. The dependent variable in Column 1 is a dummy for whether the case receives a downwards departure. The dependent variable in Column 2 is a dummy for whether any prison sentence is given, the dependent variable in Column 3 is the sentence length (in months) including sentences of zero length, and the dependent variable in Column 4 is the median sentence length. In all Columns we condition on defendant ethnicity (White, Black, Hispanic) whether the case comes up post 9-11, and interactions between the two, and three way interactions between a post 9/11 dummy, a dummy for the 2001 NE period, and ethnicity. Throughout the following additional controls are included: 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. At the foot of each Column we report the estimate of the common impact, the difference between the sentenced post-9/11 x 2001 interaction and the sentenced post-9/11 dummy, its standard error and confidence interval.

Table A10: The Patriot Act and the Common Impact on Any Prison Sentence

Dependent Variable: Any Sentence
Standard errors in parentheses clustered by ethnicity-district

	(1) All Offenses	(2) Non-Patriot Act Offenses	(3) Patriot Act Offenses (Money Laundering and Immigration)
Sentenced post 9-11 and Pre-Patriot Act	0.010	0.009	0.025
	(0.009)	(0.010)	(0.036)
Sentenced post 9-11 and Pre-Patriot Act*Black	0.004	0.002	0.087
	(0.012)	(0.012)	(0.054)
Sentenced post 9-11 and Pre-Patriot Act*Hispanic	0.002	-0.002	-0.010
	(0.011)	(0.011)	(0.037)
Sentenced post 9-11 and Post-Patriot Act	0.017***	0.015**	0.080***
	(0.006)	(0.006)	(0.027)
Sentenced post 9-11 and Post-Patriot Act*Black	-0.014*	-0.014*	0.002
	(0.008)	(0.008)	(0.040)
Sentenced post 9-11 and Post-Patriot Act*Hispanic	-0.004	-0.003	-0.054*
	(0.007)	(0.008)	(0.027)
Black	0.030***	0.029***	0.006
	(0.006)	(0.006)	(0.030)
Hispanic	0.057***	0.062***	0.044**
	(0.007)	(0.007)	(0.019)
Controls	Offender characteristics, defense counsel type, offense type dummies, guideline cell dummies, and Federal district dummies.		
p-value: [post1-Hispanic = post2-Hispanic]	0.531	0.952	0.195
p-value: [post2-Black = post2-Hispanic]	0.112	0.135	0.060
Adjusted R-squared	0.453	0.482	0.262
Observations	40,228	32,930	7,298

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. OLS regression estimates are shown in all Columns. The dependent variable is a dummy for whether the case receives a downwards departure. Standard errors are reported in parentheses, where these are clustered by ethnicity-district. The Natural experiment sample of 40,228 Federal cases is used (those that come up for sentencing in a six month window either side of 9/11/2001). For those defendants sentenced after 9/11/2001, the last offense was committed prior to 9/11/2001, and if sentenced before 9/11/2001, the last offense was committed at least 180 days prior to 9/11/2001. The Patriot Act was introduced on 26/10/2001, 45 days after 9/11. Hence we split the treated group of defendants into those sentenced post 9-11 and pre the Patriot Act coming into force, and those sentenced post 9-11 and post the Patriot Act coming into force. In all Columns we condition on defendant ethnicity (White, Black, Hispanic) and these two treatment dummies, and the following additional controls are included: 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. In Column 2 we restrict the sample to cases involving offenses that are not money laundering or immigration, and in Column 3 we restrict the sample only to money laundering and immigration offenses. The p-value at the foot of each Column is on the null that the coefficients on the post 9-11 x Black and post 9-11 x Hispanic dummy interactions are equal against a two sided alternative for each ethnic group.