Do You Receive a Lighter Prison Sentence Because You are a Woman or a White? An Economic Analysis of the Federal Criminal Sentencing Guidelines *

Todd Sørensen, University of California - Riverside and IZA
Supriya Sarnikar, Westfield State University
Ronald L. Oaxaca, University of Arizona and IZA

July 31, 2012

Abstract

Using data obtained from the United States Sentencing Commission's records, we examine the extent to which the Federal Criminal Sentencing Guidelines curbed judicial sentencing preferences based on gender, race, and ethnicity. Our structural utility maximization model of judicial sentencing and a new generalized nonlinear decomposition methodology allow us to conduct a counterfactual exercise examining the impact of the guidelines on sentences during our period of study. Our results indicate that, under the guidelines, and after controlling for circumstances such as the severity of the offense and past criminal history, judicial preferences strongly favor women while also disadvantaging black men. We find that in the absence of the guidelines the estimated unadjusted sentencing gaps would have diminished slightly but judicial preferences would have increased the unexplained gap. Our findings stand up to a wide variety of robustness checks.

*We gratefully acknowledge the helpful comments of WPEG Conference participants at the University of Kent, seminar participants at IZA, ERMES and University of Paris II, Syracuse University, University of Arizona Law College, Düsseldorf Institute for Competition Economics, University of California Riverside, University of Canterbury, Victoria University of Wellington, University of Otago, Melbourne Institute of Applied Economic and Social Research, Monash University, and Jonah Gelbach. Any remaining errors are the responsibility of the authors.

1 Introduction

In the 1980s, the Federal Criminal Sentencing Guidelines were established in part to reduce perceived disparities in sentences. We address the question of whether or not they impeded judicial proclivities to steer prison sentences in favor or against convicted defendants on the basis of their race/ethnicity or gender. The answer to this question is important because it identifies the limits to what a much heralded reform can accomplish in terms of attenuating gender and racial disparities in sentencing outcomes.

The full extent of racial/ethnic and gender discrimination in the criminal justice system is not a question that we propose to answer here. This issue would have to consider the entire process starting with decisions governing arrests, charges, prosecution, acquittal vs. convictions, and prison sentences. Instead, we focus on the preferences of the sentencing judge. Our conceptual framework posits a judicial random utility function over sentences that is subject to utility costs of departures from the guidelines. The motivation for this structural approach is twofold. First, our data contain mass points that make a linear model unsuitable. Specifically, 13% of all criminal sentences involve no prison time, 21% of sentences fall at guideline lower bounds and 6% occur at guideline upper bounds. Second, by developing a model which accounts for the bunching of sentences at these mass points, we are in a position to analyze how sentences would have been handed down had the guidelines not existed during our period of study.

Accordingly, we treat the sentencing outcome variable from utility maximization as a mixed discrete-continuous variable. The econometric model implied by our random utility model is what might be termed a partially uncensored ordered probit model. We then estimate our model for six different racial/ethnic and gender groups using sentencing data from a period during which the guidelines were in effect. Our key conditioning variables are the criminal severity and history scores that mechanically determine the sentence range permitted under the guidelines. We acknowledge that judges may have some ability to influence these scores. Assuming that judges will harmonize the score with their preferred sentences, our results will represent a lower bound of potential judicial bias. We then decompose sentencing gaps using a new decomposition methodology that is appropriate to the estimated econometric model of sentencing outcomes. This decomposition builds upon the generalized decomposition developed in Neumark (1988) and Oaxaca and Ransom (1988, 1994). Finally, we conduct a counter-factual analysis of sentences in the absence of the guidelines by setting the estimated utility costs to the judge of departing from the guidelines to zero.

Our analysis yields a number of interesting results. First, we find strong evidence of favoritism towards females vis-a-vis males among whites, blacks, and hispanics as well as discrimination against black males vis-a-vis white males. Second, we find evidence that these unjustified gaps in sentences would increase to some degree in the absence of the guidelines. Finally, we find evidence of larger unjustified sentencing gaps when using a measure of criminal severity that judges should be less able to influence. This is evidence that our results inform us on a lower bound of discrimination, as judges may be able to mask some of the discrimination by harmonizing the calculated severity score with their own preferred sentence. It also serves as a reminder that the effect of preferences often continues to persist even after legislative reform attempts to curb the expression of those preferences. We conduct a number of robustness checks and find that the above results generally hold.

2 Literature Review and Background

2.1 Brief History of the Federal Criminal Sentencing Guidelines

Until the mid-1980s, federal judges enjoyed wide discretion in the determination of criminal sentences. The establishment of the parole board in 1910 had decreased that discretion somewhat and made sentences partially indeterminate. The discretion vested with parole boards was supported by those who believed in rehabilitation of prison inmates, until it became apparent that parole was not achieving the rehabilitative purpose. Those who believed in rehabilitation became increasingly unhappy with the perceived disparities in sentencing of like crimes and with the perceived abuses of discretion by judges. Believers in deterrence as a goal of sentencing were also equally disenchanted by the perceived leniency of judges. The legislative stage was set to strip judges of their wide discretion and provide more determinate sentencing. See Stith and Koh (1993) for a detailed legislative history of the guidelines.

The passage of the Sentencing Reform Act (SRA) in 1984 created the United States Sentencing Commission (USSC) which was vested with the power to develop mandatory guidelines that would reduce "unwarranted" disparities in sentencing. Congress did not define what "unwarranted" meant, rather it was up to the sentencing commission to determine which disparities were unwarranted. The federal sentencing guidelines formulated by the commission essentially determined that disparities based on race, gender, age, income, number of dependents, etc., were unwarranted. The core of the guidelines is a matrix called the sentencing table, which contains a range of allowable sentences for each level of offense severity and criminal history category.[1]

Judges were allowed to depart from the mandatory guideline range only for reasons that were not adequately considered by the sentencing commission in formulating the guidelines. Judges were also required to provide a written rationale for departures that were not requested by the prosecutor. Prior to the guidelines, there was virtually no appellate review of a trial judge's sentencing decision. After the guidelines, departures could be appealed by the prosecutors or the defendants. The calculation of offense level (severity) was also made reviewable by a higher court.

[1] See www.ussc.gov/2009guid/TABCON09.htm.

Even though the guidelines removed much of the judicial discretion in sentencing, judges still enjoyed some leeway in the determination of offense levels. While the base offense level is determined by the charge of conviction, judges could adjust the base level up or down based on a number of fact driven characteristics. For example, the base offense level may be adjusted upwards by 2 to 3 points based on the victim's characteristics, and by 2 to 4 points if the defendant played an aggravating role in the crime. The calculation of offense level is generally accorded deferential treatment by higher courts while departures are scrutinized more closely.

In calculating the offense severity and imposing the final sentence, judges were often required to consider facts that were not determined at trial by a jury. This ultimately led to the guidelines failing the test of constitutionality in United States v. Booker, 125 S. Ct. 738 (2005). Booker turned the guidelines into advisory rules instead of mandatory.

The sentencing guidelines transferred a large part of sentencing discretion to prosecutors by making the sentence so strongly depend on the offense level. Because the prosecution and the defendant could more accurately predict what the sentence would be for different charges, the guidelines led to an increase in "charge bargaining"and "fact bargaining". Charge bargaining occurs when the prosecution negotiates a plea deal based on a lower charge than the initial charges at arraignment. Fact bargaining occurs when the plea deal involves stipulation of a set of facts that would result in a sentence range under the guidelines that is more desired by the parties to the plea deal. A change to Federal Rules of Criminal Procedure in 1999 made charge and fact bargaining even easier by allowing plea bargains that were conditional on the court accepting the negotiated sentence, or alternatively a sentence range or sentencing factor. The sentencing commission sought to combat the problem of increased prosecutorial discretion by enlisting probation officers to provide factual information that judges could rely upon for sentencing decisions. Judges were allowed to reject plea deals and consider all relevant conduct and base sentences on facts whether these were present in the plea deal or not. But according to a report[2] released by the USSC, overworked trial judges rarely question the facts as stipulated by the plea bargain. Probation officers also do not "disturb stipulations and in any event are rarely asked for their opinion until after the plea deal is accepted." King and O'Neill (2005).

In theory, judges do retain some discretion in offense level calculations as pointed out in Schanzenbach and Tiller (2007). If judges opt to take additional facts into account or adjust the offense level based on facts not found in the plea deal, the sentencing range can change quite dramatically, especially in drug cases. If indeed judicial discretion in offense level calculations is a major source of sentencing disparities, our results must be interpreted carefully. We argue below that in this instance, our results can be interpreted as lower bounds estimates of judicial preferences.

[2] "United States Sentencing Commission: Fifteen Years of Guidelines Sentencing" (November 2004), p. 86.

2.2 Literature Review

A number of studies have examined issues related to criminal sentencing. One strand of this literature has looked at the optimal degree of judicial discretion under various assumptions includes Miceli (2008), Bar-Gill and Gazal Ayal (2006), Shavell (2005), Reinganum (2000), and Reinganum (1988). Our focus is rather on how the sentencing guidelines affect race and gender differences in sentencing outcomes.

Since one of the primary goals of sentencing policy over the last few decades has been to avoid such disparities, it is important to properly estimate whether the guidelines actually reduced these unwarranted disparities. Empirical papers that have attempted to do this take different approaches to measuring the variation in sentencing. Waldfogel (1998) distinguishes between "good sentencing variation" and "bad variation" and asks whether the guidelines reduced the "bad variation" that occurs due to "capricious or malicious" behavior of judges (p. 304). His results, based on an analysis of federal criminal cases decided in California between 1984 and 1987, suggested that the guidelines would not be effective at reducing only the bad variation; they would reduce fairness enhancing "good variation" as well. Bushway and Gelbach (2010) model judge's preferences in the context of bail setting and find evidence that, holding other characteristics constant, judges in a number of counties discriminate against blacks.

Lacasse and Payne (1999) analyze cases that arose in the federal district courts of New York during 1981 – 1995, and find that the amount of variation attributable to judges increased after the guidelines went into effect. These results are based on a regression of the prison term on a set of indicator variables for the judge assigned to the case, a indicator variable for post-reforms cases, indicator variables for offense type and a selection term indicating the regime choice of plea or trial. In contrast, Anderson, Kling, and Stith (1999) examine the inter-judge disparity in average length of prison sentences and find that the disparity in sentences declined after the guidelines went into effect. Using data on a sample of cases that were assigned to judges deemed to be "active", the disparity in sentencing was measured as the dispersion of a random effect in a negative binomial model. To account for variation in the mean prison term that might come about from differences in the type of offenses pre- and post-guidelines, the authors used a set of weights based on the shares of offenses from 1986-87 in both periods. Abrams, Bertrand and Mullainathan (2010) also find evidence of racial disparities and variation across judges in incarcerations. While we do not possess data that allows us to identify the individual judge, our approach will allow us to identify the preference based behavior of the average judge.

Schazenbach (2005) used the political, racial and gender composition of the bench at the district level to estimate the effect of judicial characteristics on sentencing disparities. The race and gender bias inferred in Schazenbach (2005) is also based on estimated coefficients of indicator variable indicators of race and gender. Schazenbach and Tiller (2007) focus on both sources of judicial discretion (calculation of final offense level as well as departure from the guideline range) and find that judges strategically use the ability to adjust offense levels and departures to achieve the sentence outcomes that are more in line with their political beliefs. They also find that the use of departures to achieve desired sentence is influenced by the degree of political alignment between the sentencing judge and the circuit court while the use of adjustments is not so influenced.

Sorensen, Sarnikar and Oaxaca (2012) find evidence of racial and gender gaps in sentencing, as well as evidence that judges took into consideration other characteristics, such as age and marital status, that should not have been considered under the guidelines. Here, we expand that analysis to study a wider set of groups, and also address how sentencing disparities were affected by the existence of the guidelines themselves.

Most relevant to our own study is that of Mustard (2001). He used sentencing data from the USSC on cases resolved between 1991-1994 and found that blacks, males and offenders with low levels of income received substantially longer sentences. He also finds that departures from guidelines produced much of the disparity, with the largest black-white disparities occur in drug cases. Our study of more recent data is similar in approach to that of Mustard (2001). We expand on Mustard's (2001) analysis by allowing for a full interaction of the weights on characteristics between the groups in question. Additionally, our non-linear estimator allows us to study the bunching of sentences that occur at the high and low end of the guidelines, and how this varies across race and gender. Further, we are able to estimate how mean sentences would have been different if the bunching at the guidelines had not occurred. The question of the effect of the absence guidelines has to our knowledge only been addressed by Nutting (2012), who finds evidence that sentences dropped for women after Booker.

3 Data

The data used in this study are obtained from the United States Sentencing Commission's data collection efforts and pertain to cases that terminated in convictions over the period 1996-2002. The data set is available from the Federal Justice Resource Statistics Center (Bureau of Justice Statistics, 1997a, 1997b, 1998, 1999, 2000, 2001, 2002). We construct our cleaned dataset by including all observations where we know the individual to be between 18 and 65 years old at the time of sentencing, where we know the individual to a be a U.S. citizen, where gender and race of the individual are known, dropping individuals who are not black, white or hispanic, dropping individuals with sentences not of finite determined length (i.e. life sentences or the death penalty), and for whom potential control variables are non-missing.[3] The vast majority of the

[3]Our raw dataset has around 380,000 observations. Our final dataset contains around 85,000 observations. Of the 300,000 observations that are lost in the data cleaning process, around 150,000 are a consequence of often missing data on the type of defense counsel. In

sentencing cases were the result of guilty pleas (95%) as opposed to conviction by trial (jury or bench). Consequently, we focus only on the guilty plea cases.[4] Individuals appear in our dataset after being convicted of a number of different types of Federal offenses. While there are over 30 different types of offenses that appear in our data and types of offenses do vary across years, a small number of offenses dominate. Specifically, drug trafficking accounts for around 40% to 45% of offenses, gun charges around 10%, and fraud around another 15%.[5]

Table 1 presents the means of our outcome variable (length of sentence in months) and control key variables. These two control variables for the sentencing guide lines are captured by the severity of the final/current offense (XFOLSOR) and criminal history (CRIMHIS). Both variables are constructed from measures set according to fixed formulas established by the U.S. Sentencing Commission. To calculate the offense level, the case is assigned a base level offense and then adjusted for various aggravating circumstances such as the use of a firearm in the crime or obstruction of justice, or for mitigating circumstances such as acceptance of responsibility. The criminal history measure is a function of both the length of prior imprisonments and how recently these sentences were given.[6]

We see that women in each racial/ethnic group receive lower sentences than their male counterparts. We can see also that for all three groups of women, both criminal severity and history are lower than for men in their group. Along racial lines, the most striking difference is between white males and black males, the latter receiving an average sentence almost twice as long as the former. As with the cross gender comparisons, we also see that the group receiving the higher sentence has both higher average severity and history. This evidence suggests that at least *some* of the sentence gap is justified. However, further analysis is needed in order to determine *how much* of the sentence differential is justified by differences in circumstances (severity and history).

While clearly there are instances in which the final offense level (XFOLSOR) is determined in part by the sentencing judge, estimation of the sentencing model does not take account of the possibility that the criminal severity score may be endogenous. The potential bias that might arise from endogeneity in the calculation of the final offense level would be two-fold. First, any judicial preference based on gender and race would influence the calculated final offense level so as

---

the appendix, we confirm that our main results are robust to the inclusion or exclusion of these observations. Around 66,000 observations are lost on account of being non-citizens or of citizenship status being missing. Our age, gender, race, and ethnicity restrictions lose another 30,000 observations.

[4]The appendix also presents results confirming the robustness of our results to the inclusion of sentences handed down after a convication by trial.

[5]These figures were calculated using the OFFTYPE and OFFTYPE2 variables for the years 1996-1998 and 1999-2002, respectively.

[6]For details on the construction of these variables, please see the following documents on the USSC's website:

http://www.ussc.gov/training/sent_ex_rob.pdf
http://www.ussc.gov/training/material.htm

to minimize the apparent racial and gender-based sentencing disparities. This should lead to a downward bias or underestimation of the unexplained racial and gender based sentencing gaps. Second, the presence of utility costs for departures from the guidelines would be expected to encourage judges to "harmonize" the calculated criminal severity index scores so as to make it less likely that a judge would appear to be departing from the guidelines. This would lead to underestimates of the frequency of departures and in the utility costs of departures. Left untreated, the potential bias will lead to underestimation of race and gender based judicial preferences, departures from the guidelines, and utility costs of departure. So any manifestation of judicial bias, guideline departures, and utility costs of departures from our estimated sentencing model would indicate the actual presence of these factors despite the downward bias. Later, we report on evidence of this *harmonization* effect by using a more preliminary severity measure (BASEHI). Judges should have less of an influence over this measure, and we indeed see evidence of more discrimination when using it.

4 Model

Our model of judicial sentencing preferences holds that the judge seeks to maximize their utility over the ideal sentence for a convicted defendant, subject to utility costs from departures from the sentencing guidelines. We specify a quadratic utility function

$$U_i = \frac{-1}{2}\left(S_i - S_i^{*}\right)^2 - \theta_h(S_i - G_i^h)(D_i^+) - \theta_l\left(G_i^l - S_i\right)(D_i^-)$$

where for the *ith* convicted defendant, U_i is the sentencing judge's utility, S_i is the sentence awarded, S_i^* is the ideal sentence in the absence of costs from departures from the sentencing guidelines (sentencing bliss point), G_i^h is the maximum sentence specified by the guidelines, G_i^l is the minimum sentence specified by the guidelines, $0 \leq G_i^l \leq G_i^h$, D_i^+ and D_i^- are indicator variables for upward and downward departures from the guidelines and are defined by $D_i^+ = 1\left[S_i > G_i^h\right]$ and $D_i^- = 1\left[S_i < G_i^l\right]$. The parameter restrictions are $\theta_h, \theta_l > 0$.

Utility maximization implies the FOC:

$$\frac{\partial U_i}{\partial S_i} = -\left(S_i - S_i^*\right) - \theta_h D_i^+ + \theta_l D_i^- = 0$$

which yields the sentencing function

$$\dot{S}_i = S_i^* - \theta_h\ D_i^+ + \theta_l D_i^-,$$

where $\dot{S}_i$ is the constrained utility maximizing sentence. Note that for a judge for whom $D_i^+ = 1$ for a given defendant, it is the case that $\dot{S}_i - S_i^* = -\theta_h < 0$. In other words, the utility maximizing sentence is below the ideal sentence. The judge would depart upwards from the guidelines but not as much as would have been preferred in the absence of the guidelines. Similarly, for a judge for whom

$D_i^- = 1$ for a given defendant, it is the case that $\dot{S}_i - S_i^* = \theta_l > 0$. In other words, the constrained utility maximizing sentence is above the ideal sentence. The judge would depart downwards from the guidelines but not as much as would have been preferred in the absence of the guidelines. Actual sentences deviate from ideal sentences whenever the guidelines are binding.

Note that some departures from the guidelines can only be initiated by the prosecution. These are known as 5K1.1 departures after the policy guideline that governs them. Because these departures are not initiated by the sentencing judge, we explore how our results differ based on the inclusion of these cases. Accepting prosecution's 5K1.1 motion and departing from the guidelines may be less likely to lead to an appeal, but these departures also impose a utility cost on the judge in terms of attracting additional scrutiny from the sentencing commission and outside researchers. Our empirical results present evidence of judicial bias both including and excluding these departures.

We assume that the ideal sentence is specified by the stochastic function

$$S_i^* = X_i'\beta + \varepsilon_i \tag{1}$$

where X_i' is a vector of the defendant's characteristics and facts of the case that determines the judge's preferences for the ideal sentence, β is a vector of parameters, and ε_i represents random utility and is distributed *i.i.d.* $N(0, \sigma_\varepsilon^2)$.

Given the threshold nature of the guidelines, the actual sentence awarded is based on a utility maximization problem which spans 6 regions:

Discrete	**Continuous**
$S_i = 0$	$0 < S_i < G_i^l$
$S_i = G_i^l$	$G_i^l < S_i < G_i^h$
$S_i = G_i^h$	$G_i^h < S_i$

Cox and Oaxaca (1982) derive a simple Tobit model of the median legislator's preference for a minimum wage based on utility maximization. Generalizing their utility model, we obtain a model that we term the "partially uncensored ordered probit model", which allows for mass points at 0, G_i^l and G_i^h, as well as for a continuously distributed variable in the other three regions. Our utility maximization behavioral assumption and specification of S_i^* yields the likelihood function and expected values for this model, allowing us to use Full Information Maximum Likelihood (FIML) to estimate the β and θ parameters and compute the expected length of sentences needed for the decompositions that we describe next. Details on our estimator are found in Appendix C.

4.1 Decomposition Methodology

To examine how much of the gender/race differences in sentences can be ascribed to leniency toward one group or another, we estimate gender/race preferences in criminal sentencing outcomes by applying empirical methods developed in the labor economics literature. These methods have the advantage of decomposing group differences in sentencing outcomes into three different components: one due to differences in the observable circumstances of the convictions and the other two pertaining to judicial preferences for each group in a binary comparison. Conventional decompositions adopt the outcome model of one group as the norm and predict the outcome of the other group from this norm, e.g. Oaxaca (1973). This approach attributes all of the unexplained outcome gap between two groups as the result of the second group's outcome deviating from that predicted on the basis of the outcome model for the first group. A generalized decomposition methodology exists that permits one to apportion the unexplained outcome gap to a positive preference for one group and a negative preference for the other group, e.g. Neumark (1988), Oaxaca and Ransom (1988, 1994). A natural norm for the generalized decomposition is the estimated model obtained from pooling the groups being compared. In the present case of the partially uncensored ordered probit model, we estimate the model with a pooled sample of all observations and also separately for each of the six demographic groups. The predicted mean sentences are obtained from

$$\hat{S}_j^0 = \frac{\sum_{i=1}^{N_j} \hat{S}_{j_i}^0}{N_j}, \text{ (pooled model)}$$

$$\hat{S}_j = \frac{\sum_{i=1}^{N_f} \hat{S}_{j_i}}{N_j}, \text{ (own model)}$$

where j is a place holder for the group being examined (male, female and black, white, or hispanic).

Consider a decomposition of the mean sentence between two groups, k and l. The decomposition of observed sample mean sentences proceeds according to

$$\begin{aligned} \bar{S}_k - \bar{S}_l &= \left(\hat{S}_k - \hat{S}_l\right) + \hat{\delta}_{kl} \\ &= \left(\hat{S}_k - \hat{S}_k^0\right) + \left(\hat{S}_l^0 - \hat{S}_l\right) + \left(\hat{S}_k^0 - \hat{S}_l^0\right) + \hat{\delta}_{kl} \end{aligned} \tag{2}$$

where $\bar{S}_j = \frac{\sum_{i=1}^{N_j} S_{j_i}}{N_j}$ for both k and l. The term $\hat{\delta}_{kl}$ is the difference between the sample mean sentencing gap $\bar{S}_k - \bar{S}_l$ and the predicted mean sentencing gap $(\hat{S}_k - \hat{S}_l) = \left(\hat{S}_k - \hat{S}_k^0\right) + \left(\hat{S}_l^0 - \hat{S}_l\right) + \left(\hat{S}_k^0 - \hat{S}_l^0\right)$. The term $(\hat{S}_k - \hat{S}_k^0)$ estimates judges' sentencing preferences toward group k (when compared to the population as a whole), the term $\left(\hat{S}_l^0 - \hat{S}_l\right)$ measures sentencing preferences toward group l (when compared with the population as a whole), and the term

$\left(\hat{S}_k^0 - \hat{S}_l^0\right)$ estimates the portion of the predicted sentencing gap attributable to differences in the case circumstances (between groups k and l).

4.2 Absence of Guidelines

In order to determine the extent to which the U.S. Federal sentencing guidelines curbed judicial sentencing preferences based on gender and race, one needs to construct a counterfactual. The judicial utility maximization model corresponding to the counterfactual absence of the sentencing guidelines (ignoring statutory sentencing limits) would simply be

$$U_i = \frac{-1}{2}\left(S_i - S_i^*\right)^2$$

with the FOC given by

$$U'(S_i) \leq 0$$
$$S_i \cdot U' = 0.$$

By setting our estimates of the θ terms to 0, we are able to construct the expected sentence length along the lines of a standard Tobit model and decompose the expected sentencing gap between two groups as shown above.

5 Empirical Results

5.1 Preliminary Evidence

Our later analysis will control for the bunching in sentences length around the upper and lower end of the guidelines shown in Figure 1, and also allow flexibility across groups in the sentence determination process. To begin, we run a simple OLS regression pooling all of our observations together and including indicator variables for each of the six groups for which we report mean sentences in Table 1. In this estimation we control for a number of factors that may affect sentences.[7] These controls include demographic variables such as education, material status, number of dependents and age. The also include an indicator for whether the individual had a private attorney, indicator variables for the criminal history category, a cubic of the criminal severity score, and fixed effects for the circuit and year in which the sentence was handed down. These variables are described in Table 2 and summary statistics are reported for each demographic group in the appendix.

[7] Many of these control factors, such as education and access to a private defense attorney may differ across race and gender as well. To the extent that differences in these factors result in an unequal distribution of resources across race and gender, our results represent a lower bound of unjustified differences in sentence length as a whole. However, our intention is to estimate how much of the sentence gap can be attributed to judicial bias. With the possible exception of the criminal severity score (discussed at more length elsewhere), the sentencing judge should have no influence over our control factors. Thus, they are appropriate to condition on if our intention is to estimate judicial bias in sentencing.

Table 3 presents the results of our linear estimation on two different samples: 1) all sentences in our final dataset (left column) 2) only sentences in our final that fell within the guideline range (right column). In the left column, we see strong evidence that sentences differ across groups, even when conditioning on important factors. The coefficient on the indicator variable for all three female groups is negative and significant, indicating that females receive sentences that are 4.45 to 5.44 months shorter than white males, on average.[8] The positive and significant coefficient for black males suggests that, even after conditioning on severity and history, African American males receive higher sentences than do white males. We find no significant differences between white and hispanic males. When we limit our analysis to cases *within* the guidelines, we see a somewhat different picture. Here, we find evidence that (at the 5% level) white women receive *higher* sentences than white men, and other women receiving sentences that are at least not lower than those of white men. We no longer find evidence of black men receiving higher sentences. These findings paint us two different pictures: when looking within the guidelines, where the law reduces judicial discretion and sentences based on the judges preferences would arguably be easier to observe, we see little evidence of the type of preferences that appear when we look across all sentences. In summary, Table 3 both provides us with some more preliminary evidence that judicial preferences may affect average sentence length across groups and also shows the importance of considering the role of the guidelines (and departures from them) in our more rigorous analysis.[9]

5.2 Utility Maximization Model

5.2.1 Estimation Results

Table 4 presents results from our utility maximization based model. The table presents seven sets of estimates: in the left column are the estimates from a pooled sample, and the next six columns present results for each of our groups. We see that sentence length is monotonically increasing across our criminal history category for all groups (though each coefficient is not always significantly different from the previous). While each term in the criminal severity score (XFOLSOR) cubic is not always significant, the three terms are always jointly significant at the 99.99% level. Additionally, the mean and median marginal effects of an increase in severity is positive for each group and are on the order of between around 1.50 and 4.28 months of additional sentence length. As shown by in Sorensen, Sarnikar and Oaxaca (2012), significant coefficients on many of the demographic characteristics suggest that judges are taking into consideration factors that were not generally allowed to influence sentence length under the guidelines.[10]

[8]We cannot reject an F-test of the equality of these three female coefficients, suggesting that all three of these groups of females receive the same amount of additional leniency when compared to white males.

[9]Full results from these linear estimates are available upon request from the authors.

[10]In addition to race, sex and national origin, the guidelines also exclude factors such as socioeconomic status, family ties and responsibilities, and (with only limited exceptions) age

The bottom panel of Table 4 reports estimates of the θ terms. As we discussed in the previous section, these terms represent the utility cost to judges of departing from the guidelines. They are identified by the degree of bunching that we observe at the top and bottom of the guidelines. Their relative values inform as to the relative cost to judges of departing from the guidelines for each group and can inform us about sentence gaps between groups. Their absolute values will help to inform us as to how sentences would have been handed down differentially had the guidelines not been in place. Our estimates suggest that judges find it more costly to depart downwards than upwards: our pooled model finds that a judge would rather depart 46.17 months from their ideal sentence than depart one month from the lower end of the guidelines and would rather depart 25.99 months from their ideal sentence than depart one month from the upper end of the guidelines. The higher utility cost of departing downward than upward is consistent with the guidelines being established in response to concerns about excessive leniency by judges. We see that the three lowest downward departure cost estimates are for the three female groups, while the highest estimate is for black males, all consistent with the observed sentencing gap. Our group-wise estimates of the upward departure costs produce counter intuitive results. However, upward departures occur at a much lower frequency and thus should not drive our decomposition results.

5.2.2 Decomposition and Counterfactual Analysis

With the above parameter estimates, we are able to compute two average predicted sentence lengths for each group: 1) using the parameters from an estimation of our model using only individuals from that group, and 2) using parameters from an estimation of our model pooling all data. The former generates a baseline prediction, the latter tells us how sentences would have differed had this group been treated like all other individuals.[11] By comparing predictions by own and pooled parameter estimates, we are able to decompose the sentencing gap, as discussed in the previous section.

Table 5 presents our decomposition estimates. The top panel decomposes the gender gap between our three racial/ethnic groups, while the bottom panel decomposes four racial/ethnic gaps. For each group wise comparison, we present a decomposition of the gap under the guidelines as well as a decomposition of the gap under the assumption that the guidelines had been removed entirely (in other words, the utility cost of departure had been set to 0). For each decomposition, the first column presents the observed gap between the two groups and the second presents the predicted gap in our model (which is what

and education.

[11]In Appendix Table A7, we report our expected sentences by group and parameters used. The table also shows the likelihood of being in each of six regimes (1) 0 sentence, 2) strictly positive but below lower end of guidelines, 3) at lower end of guidelines, 4) within guidelines, 5) at upper end of guidelines, 6) above upper end of guidelines, the expected value of the sentence for individuals that fell into each regime, as well as an overall expected value. Please note that the overall expected value is not necessarily equal to a weighted average of the expected values of of individuals within each regime, due to non-linearities.

we decompose); no observed gap exists in the absence of the guidelines. The third column shows how much of the predicted gap can be explained by different circumstances facing the two groups in question, while the fourth column gives the remaining gap that is unexplained. The unexplained gap represents an estimate of the portion of the total predicted gap which may be attributed to bias stemming from judicial preferences. Of course, as we have noted above, this will simply be a lower bound of this potential bias if judges influence the criminal severity score in the same direction as their preferences affect sentences. Because of our use of Oaxaca and Ransoms's (1988) generalized decomposition approach, we are able to further decompose the unexplained gap into portions attributed to favoritism or discrimination towards each group relative to the general population.

The top panel of Table 5 examines the gender sentencing gap and finds large unexplained gaps for all three groups. Of the 20.64 predicted gap between average white male and white female sentences, only 14.34 of these months can be explained by different circumstances facing the two groups. The additional 6.3 months remain unexplained and can be attributed to judicial bias. Very little of this gap (only 0.10 of a month) can be attributed to judges treating white males more harshly than the population as a whole, while a full 6.20 months can be explained by judges treating females more leniently than the population as a whole. When we conduct our counterfactual exercise of removing the utility cost to departure, we see that the predicted gap falls slightly, the explained gap falls more sharply, and thus the unexplained gap increases. The increase in the unexplained gap is due mostly to an increased level of leniency toward females. The reason for this is straightforward: in the absence of the guidelines, sentences fall for both white males and white females, but by more for white females on account of judicial preferences. The gender gap among blacks is similar, though slightly larger than for whites on account of a larger unexplained gap for black males versus the general population (which will become even clearer once we view the bottom panel of the table). The hispanic gender gap is very similar to the white gender gap. In all cases, removing the guidelines leads to larger unexplained gaps between male and female sentences on account of more leniency towards females.

In the bottom panel of Table 5, we examine racial gaps. Here, the observed gaps and predicted gaps represent the sentence for the white group minus the sentence for the minority group. As average sentences for whites in our data are always lower than those for their minority counterparts, this difference is negative. In contrast to the gender gaps, most of the racial gap can be explained by different characteristics, such as criminal severity and criminal history, faced by the individual. For example, 28.98 months of the 31.56 of the black white gap among males can be explained by differences in circumstances. The unexplained gap of 2.58 months between black and white males is the largest gap along racial lines that we uncover in our analysis. This is mostly attributed to bias against black males rather than bias in favor of white males. Along the other three lines along which we decompose racial sentencing gaps, we find no evidence of unexplained gaps greater than one month. It should be noted that the female racial gaps include large terms in absolute value for each sub-part of our unexplained gap. This simply means that both minority and white women are treated different than the average individual (i.e. including males), but that this treatment nets out to no significant difference between the two groups, as was alluded to in our earlier linear analysis.

We note that our analysis does not conclude that large amounts of racial bias do not exist in the criminal justice system. In fact, institutional factors such as the the crack-powder cocaine disparities in the guidelines that were in place at the time could be regarded as inherently biased against African Americans. However, this is not a case of *judicial bias*. In summary, our main decomposition results find that 1) judicial preferences lead to lower sentences for otherwise equivalent males and females, 2) these preferences would produced even greater gender sentencing gaps in the absence of the utility cost of departing from the guidelines and 3) small unexplained racial sentencing gaps exist for most groups.

As we have previously noted, judges may influence the criminal severity score. If this were the case, one of our key conditioning variables would be endogenously determined. If a judge's preferences affect sentences and the severity score in the same direction, some of the judicial sentencing bias would be hidden. We acknowledge this concern, and the potential that our results represent only a lower bound of the potential discrimination.

While we believe that simply estimating a lower bound of the sentence differential is informative both to other researchers and policy makers, we pursue an alternative, plausibly more exogenous, measure of criminal severity. In Table 6 we present results using the preliminary score, which the judge should have less an ability to influence. We hope that examining these additional results will help to inform us about potential biases in the estimation of the unexplained gap. Our results in Table 6 produce almost uniformly higher estimates of the total unexplained gap (the exceptions being cases where we did not find large levels of discrimination in our original analysis, along with some of the counterfactual analysis). These results indicate that judges' ability to impact the criminal severity score may indeed mask some portion of the discrimination that is taking place and underline our earlier assertion that our main results should be considered a lower bound of the role of judicial preferences.

5.2.3 Further Robustness Checks

In Appendix B we report decomposition results from a number of robustness checks. Tables B1 and B2 look at sentencing years exclusively before 1999 and after 1999, respectively. We break our sample up to test for robustness to F.R.Crim.P that made charge and fact bargaining easier. We see no systematic changes in our estimates of the unexplained gap: results are quantitatively similar and the gap shrinks for some groups and grows for others between these two periods.

In Table B3 we calculate our decomposition after reconstructing our dataset by not dropping observations for which information on the type of defense counsel is missing; we then omit this variable from our estimation. This greatly expands the size of our pooled regression to over 191,000 observations. We see that our main results are robust to the inclusion or exclusion of these observations. In these estimations, convergence was achieved for all groups with the exception of white females, for whom it was forced after 200 iterations for white females.

Our baseline results include both circuit and year indicator variables to control for changes in sentence length by group both over time and different average sentences by group across circuits. In Table B4 we have attempted to control for changes over time in each circuit by including circuit specific time trends. Estimating these additional parameters was non-trivial in our non-linear model, and convergence was forced after 200 iterations in the estimation of the model for white females. While the non-convergence may limit the reliability of these estimates, we do in fact find very little difference between these estimates and our main estimates that do not include circuit specific time trends.

Our final two robustness checks are reported in Tables B5 and B6. In these two tables, we change our sample construction. In Table B5 we have dropped all downward departures given on account of substantial assistance to the prosecution. Here, we find that for groups where there was a noteworthy unexplained gap in our main results, the gap has fallen significantly (i.e. for all gender comparisons, and for the comparison between black and white males). This provides some evidence that bias in the system may also occur because different groups are offered the opportunity to bargain down their sentences with prosecutors at different rates, or because co-operators from different groups are treated differently by judges. Finally, in Table B6 we include observations for which conviction was reached by trial. Here, we find the results to be similar to the results for our main sample.

6 Conclusion

In determining how much the U.S. Federal Sentencing guidelines impinged on judicial sentencing preferences along gender, race, and ethnic lines, it is necessary to determine how much of the sentencing gaps can be explained by circumstances of the sentencing cases other than gender, race and ethnicity. Unlike previous studies, our study employs a decomposition strategy that allows us to separate observed differences in sentencing into two different components – one attributable to differences in case circumstances, and the second attributable to differences in attitudes of sentencing judges towards defendants. The latter is further subdivided into judicial preferences toward each group within a binary comparison. Judicial preferences are captured by a quadratic utility function of the deviation between actual and ideal sentences. Depending on judicial preferences, the presence of sentencing guidelines along with the implicit costs of departures from the guidelines stand in the way of attaining the bliss point of the judges' ideal sentences.

Our econometric model is estimated by FIML and follows directly from the stochastic specification of the utility maximization model in which sentences can depart from the guidelines and in which sentencing mass points exist at 0, at the lower guidelines $\left(G_i^l\right)$ and at the upper guidelines $\left(G_i^h\right)$. Separate estimation of a pooled sample as well as one for each of our six demographic groups permit us to use a unique decomposition methodology to sort out preferences from circumstances. Such decomposition provides a better insight into the decision-making process of sentencing judges. Knowing whether judges consider extralegal circumstances in their decision making is important, but knowing how they consider extralegal circumstances is useful to policy makers in deciding how to reform sentencing guidelines to ensure equal treatment. This study not only examines whether judges consider extralegal circumstances but if they do, it asks whether they attach the same weights to circumstances of each of the demographic groups considered in this paper. We condition on the criminal severity score, which judges may have some ability to influence. We reason that judges will harmonize the score with their preferred sentence, thus leaving us with a lower bound estimate of the true unjustified gaps in sentencing.

Unconditionally, black males have the longest sentences (68 months) and white females have the shortest sentences (15 months). Our results show that judges do take into account extralegal factors when passing sentence. Under the guidelines, judicial preferences strongly favor women in all racial/ethnic groups. Judicial preferences disadvantage black males compared with white males whereas preferences regarding hispanic males are very small and consequently culminate in negligible unexplained sentencing gaps when compared with white males. The favorable preferences toward women are very close across racial/ethnic groups so that the unexplained racial/ethnic sentencing gaps are quite modest for women. The estimated marginal costs of upward sentencing departures are uniformly less than the estimated marginal costs of downward departures. This result is consistent with one of the motivations for the guidelines which was a desire to reduce perceived excess leniency in sentencing. With these parameter estimates in hand, we are able to calculate how sentences would have been distributed in the absence of the existence of the guidelines. Judicial preferences across all groups are predicted to rise in the counterfactual absence of the guidelines along with the implied unexplained sentencing gaps. At the same time, predicted total sentencing gaps would have attenuated which implies that the explained gaps would also have diminished.

Our evidence of larger unexplained gaps when using the preliminary severity score is consistent with judges setting sentencing scores to harmonize their preferred sentence with the guidelines. This unintended consequence is a reminder of the problems always present in reform meant to address socially undesirable outcomes driven by individuals' preferences.

One should bear in mind that our data permit us to examine only the end stage of the criminal justice system. A more comprehensive treatment would take account of the fact that before arriving at the judge for sentencing, a defendant must also pass through a jury or possible plea bargain with a prosecutor. Before reaching this stage, other groups, such as the police and the prosecution, have the potential to create bias in the criminal justice system. Even in light of the Supreme Court's decision in 2005 to make the Federal Sentencing Guidelines advisory instead of mandatory, our results may offer some guidance as to what to expect now that judges are less constrained in imposing sentences.

References

[1] Abrams, David S.; Bertrand, Marianne and Mullinathan, Sendhil (2010). "Do Judges Vary in Their Treatment of Race?". *Scholarship at Penn Law.* Paper 364.

[2] Anderson, James M., Jeffrey R. Kling; and Kate Stith (1999). "Measuring Interjudge Sentencing Disparity: Before and after the Federal Sentencing Guidelines." *Journal of Law & Economics,* 42(1), Part 2: 271 - 307.

[3] Bar-Gill, Oren and Oren Gazal Ayal (2006), "Plea Bargains only for the Guilty," *Journal of Law and Economics*, 49: 353-364.

[4] Bureau of Justice Statistics, Federal Justice Statistics Program. "Defendants sentenced under the guidelines during fiscal year 1996: SC96OUT", July1997a. Washington D.C. : The Urban Institute (distributor).

[5] Bureau of Justice Statistics, Federal Justice Statistics Program. "Defendants sentenced under the guidelines during fiscal year 1997: SC97OUT", July, 1997b. Washington D.C. : The Urban Institute (distributor).

[6] Bureau of Justice Statistics, Federal Justice Statistics Program. "Defendants sentenced under the guidelines during fiscal year 1998: SC98OUT", July, 1998. Washington D.C. : The Urban Institute (distributor).

[7] Bureau of Justice Statistics, Federal Justice Statistics Program. "Defendants sentenced under the guidelines during fiscal year 1999: SC99OUT". Washington D.C. : The Urban Institute (distributor).

[8] Bureau of Justice Statistics, Federal Justice Statistics Program. "Defendants sentenced under the guidelines during fiscal year 2000: SC00OUT". Washington D.C. : The Urban Institute (distributor).

[9] Bureau of Justice Statistics, Federal Justice Statistics Program. "Defendants sentenced under the guidelines during fiscal year 2001: SC01OUT". Washington D.C. : The Urban Institute (distributor).

[10] Bureau of Justice Statistics, Federal Justice Statistics Program. "Defendants sentenced under the guidelines during fiscal year 2002: SC02OUT". Washington D.C. : The Urban Institute (distributor).

[11] Bushway, Shawn D. and Jonah B. Gelbach (2010), "Testing for Racial Discrimination in Bail Setting Using Nonparametric Estimation of a Parametric Model", mimeo.

[12] Cox, James and Ronald L. Oaxaca (1982)."The Political Economy of Minimum Wage Legislation", *Economic Inquiry*, Vol.20, No.4 (October): 533-555.

[15] King, Nancy and Michael O'Neill (2005), "Appeal Waivers and Future of Sentencing Policy", *Duke Law Journal*, 55(2):209 – 262.

[16] Lacasse, Chantal and A. Abigail Payne (1999). "Federal Sentencing Guidelines and Mandatory Minimum Sentences: Do Defendants Bargain in the Shadow of the Judge?" *Journal of Law and Economics*, 42(1), Part 2: 245-269.

[18] Miceli, Thomas, J. (2008). "Criminal sentencing Guidelines and Judicial Discretion," *Contemporary Economic Policy*, 26(2): 207-215.

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

[20] Neumark, David (1988). "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *Journal of Human Resources* 23: 279-295.

[21] Nutting, Andrew (2012a). "The Booker Decision and Discrimination in Federal Criminal Sentences." *Economic Inquiry*.3

[22] Oaxaca, Ronald L. (1973)."Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review,* 14(3): 693-709.

[23] Oaxaca, Ronald L. and Michael R. Ransom. (1988). "Searching for the Effect of Unionism on the Wages of Union and Nonunion Workers," *Journal of Labor Research,* 9: 139-148.

[24] Oaxaca, Ronald L. and Ransom, Michael R. (1994)"On Discrimination and the Decomposition of Wage Differentials", *Journal of Econometrics*, 61(1), (March): 5-21.

[25] Reinganum, Jennifer (1988), "Plea Bargaining and Prosecutorial Discretion," *American Economic Review*, 78: 713-728.

[26] ________(2000). "Sentencing Guidelines, Judicial Discretion, and Plea Bargaining," *RAND Journal of Economics*, 31(1): 62-81.

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

[28] Schazenbach, M., and Emerson Tiller (2007). "Strategic Judging Under the United States sentencing Guidelines: Positive Political Theory and Evidence," *Journal of Law, Economics and Organization*, 23: 24-56.

[29] Sorensen, Todd; Sarnikar, Surpriya and Ronald L. Oaxaca (2012). "Race and Gender Differences under Federal Sentencing Guidelines, " *The American Economic Review*, 102(3): 256-60.

[30] Shavell, Steven (2005). "Optimal Discretion in the Application of Rules," *The Harvard John M. Olin Discussion Paper* No. 509.

[31] Stith, Kate, and Steve Y. Koh (1993). "The Politics of Sentencing Reform: The Legislative History of the Federal Sentencing Guidelines," *Wake Forest Law Review*, 28: 223-290.

[32] USSC (2004), "United States Sentencing Commission: Fifteen Years of Guidelines Sentencing" (November), www.ussc.gov/15_year/15year.htm.

[33] Waldfogel, Joel (1998). "Does inter-Judge Disparity Justify Empirically Based sentencing Guidelines?" *International Review of Law and Economics*,18: 293-304.

Table 1: Means of Key Variables by Group

	Sentence in Months	Crime Severity Score	Criminal History Score
White M	36.03	17.06	2.26
White F	15.23	13.91	1.59
Black M	68.06	21.05	3.10
Black F	19.33	13.90	1.76
Hispanic M	45.08	19.46	2.19
Hispanic F	20.27	16.45	1.42

Table 2: Variable Definitions

Variable	**Definition**
HISCHOOL	Indicator for High School Degree as Ed. Level
GED	Indicator for GED as Ed. Level
SOMECOLL	Indicator for Some College as Ed. Level
COLLGRAD	Indicator for College Degree Degree as Ed. Level
MARRD	Indicator for Married
NUMDEPEN	Number of Dependents
DEFENSEP	Indicator for Private Defense Attorney
CRIMHIS2	Indicator for Criminal History Category 2
CRIMHIS3	Indicator for Criminal History Category 3
CRIMHIS4	Indicator for Criminal History Category 4
CRIMHIS5	Indicator for Criminal History Category 5
CRIMHIS6	Indicator for Criminal History Category 6
XFOLSOR	Final Criminal Severity Score
XFOLSOR2	Final Criminal Severity Score Squared
XFOLSOR3	Final Criminal Severity Score Cubed
AGE	Age of Defendent
AGE2	Age of Defendent Squared
CIRC2	Indicator for 2nd Circuit Court
CIRC3	Indicator for 3rd Circuit Court
CIRC4	Indicator for 4th Circuit Court
CIRC5	Indicator for 5th Circuit Court
CIRC6	Indicator for 6th Circuit Court
CIRC7	Indicator for 7th Circuit Court
CIRC8	Indicator for 8th Circuit Court
CIRC9	Indicator for 9th Circuit Court
CIRC10	Indicator for 10th Circuit Court
CIRC11	Indicator for 11th Circuit Court
y1	Indicator for Sentencing Year 1996
y2	Indicator for Sentencing Year 1997
y3	Indicator for Sentencing Year 1998
y4	Indicator for Sentencing Year 1999
y5	Indicator for Sentencing Year 2000
y6	Indicator for Sentencing Year 2001

Table 3: Linear Model with Group Indicator Variables

	All Obs	**Within Guidelines**
Hispanic Females	-5.4429	1.3298
	(0.61767)	(0.68306)
White Females	-4.4500	1.3052
	(0.41088)	(0.44513)
Black Females	-4.7431	0.5118
	(0.48825)	(0.50759)
Hispanic Males	-0.3781	-0.5397
	(0.35021)	(0.32541)
Black Males	3.2587	0.0311
	(0.29924)	(0.26013)
N	84991	13882
R-Squared	.7	.97

Table 4: Full Estimation Results

	Pooled	**White M**	**White F**	**Black M**	**Black F**	**Hisp M**	**Hisp F**
HISCHOOL	0.7231	-1.8848	1.4325	2.5318	0.1439	2.4633	3.8544
	(0.87510)	(1.49675)	(2.13054)	(1.73175)	(2.41860)	(1.95316)	(3.23004)
GED	2.3969	1.7483	2.8676	5.4512	-4.2865	1.3104	4.0058
	(0.98380)	(1.62300)	(2.35632)	(1.99971)	(3.05351)	(2.25610)	(3.69199)
SOMECOLL	-1.8167	-4.1976	2.0862	0.4390	-2.0326	-2.5866	0.9255
	(0.93276)	(1.55651)	(2.19159)	(1.91337)	(2.52410)	(2.17799)	(3.47367)
COLLGRAD	-0.8772	-3.2412	0.8066	-1.5285	3.2351	-9.3379	-1.6945
	(1.13034)	(1.70073)	(2.62498)	(2.90496)	(3.42297)	(3.25630)	(5.09232)
MARRD	-2.5631	-4.2096	0.4035	-5.4476	-0.2441	-1.9958	3.5534
	(0.49633)	(0.79454)	(1.05401)	(1.13806)	(1.57376)	(1.07511)	(1.52147)
NUMDEPEN	-0.5528	-1.0584	-0.3507	-0.8121	-0.4739	-0.2595	-1.2909
	(0.12262)	(0.21908)	(0.31904)	(0.24613)	(0.37136)	(0.27245)	(0.42694)
DEFENSEP	-5.1702	-9.0976	-0.3306	-3.8704	1.3808	-2.9888	-2.9078
	(0.43586)	(0.65131)	(0.91288)	(1.05975)	(1.47360)	(1.02126)	(1.59305)
CRIMHIS2	10.0228	6.1166	5.0139	10.0344	6.5184	13.1481	8.4358
	(0.59634)	(0.91109)	(1.35247)	(1.40320)	(1.79645)	(1.31058)	(2.09155)
CRIMHIS3	21.2886	16.5969	10.9913	22.0642	17.6311	20.6030	14.4331
	(0.54922)	(0.88418)	(1.34113)	(1.21565)	(1.76272)	(1.16799)	(2.27019)
CRIMHIS4	34.3482	28.6257	17.7099	34.6574	27.4807	30.3712	24.1719
	(0.70427)	(1.14631)	(2.23191)	(1.40276)	(2.76332)	(1.59627)	(4.07980)
CRIMHIS5	42.4254	36.2204	28.6885	42.4530	36.7768	36.2836	28.0956
	(0.88046)	(1.44272)	(2.93732)	(1.65366)	(3.77830)	(2.13532)	(5.89810)
CRIMHIS6	58.3742	54.3775	32.8978	59.4019	44.0300	46.7384	23.3820
	(0.66424)	(1.06350)	(2.37537)	(1.31305)	(2.60682)	(1.73144)	(5.39079)
XFOLSOR	-3.5608	-3.8486	-3.5457	0.4254	-3.6458	-2.0885	-5.4367
	(0.23240)	(0.36143)	(0.48820)	(0.55475)	(0.68098)	(0.57868)	(0.89499)
XFOLSOR2	0.2634	0.2755	0.1871	0.1115	0.1869	0.1547	0.3297
	(0.01251)	(0.02001)	(0.02939)	(0.02844)	(0.04005)	(0.02921)	(0.05070)
XFOLSOR3	-0.0017	-0.0018	-0.0004	0.0000	-0.0001	0.0003	-0.0033
	(0.00020)	(0.00033)	(0.00053)	(0.00044)	(0.00070)	(0.00045)	(0.00087)
AGE	-0.3085	0.5652	0.5376	-0.8567	-0.0529	0.1636	0.6851
	(0.11531)	(0.17487)	(0.25125)	(0.29815)	(0.39620)	(0.27610)	(0.42393)
AGE2	0.0030	-0.0063	-0.0065	0.0081	-0.0013	-0.0031	-0.0087
	(0.00150)	(0.00218)	(0.00327)	(0.00406)	(0.00541)	(0.00372)	(0.00585)
Constant	-11.5218	-30.3135	-32.6401	-48.3520	-16.5796	-7.7837	-6.2852
	(2.69727)	(4.34102)	(5.99068)	(6.99587)	(9.15531)	(6.25278)	(9.24506)
σ	51.4021	49.8442	32.0804	60.1485	36.7023	46.0086	31.6795
	(0.19005)	(0.30082)	(0.44183)	(0.40490)	(0.61938)	(0.41407)	(0.66746)
θ_L	46.1740	43.4514	32.6124	52.0206	40.3907	44.4158	34.0476
	(0.32198)	(0.49392)	(0.79021)	(0.69335)	(1.13178)	(0.75782)	(1.28580)
θ_H	24.9936	25.5641	17.5070	27.8359	20.3925	22.6602	15.5521
	(0.35950)	(0.56074)	(1.07957)	(0.71705)	(1.46342)	(0.84864)	(1.83498)
N	84991	33232	7488	22834	5167	13253	3017

[a] Year and circuit fixed effects also included in all esitmations. Full results available opon request from the authors. Standard Errors in Parentheses

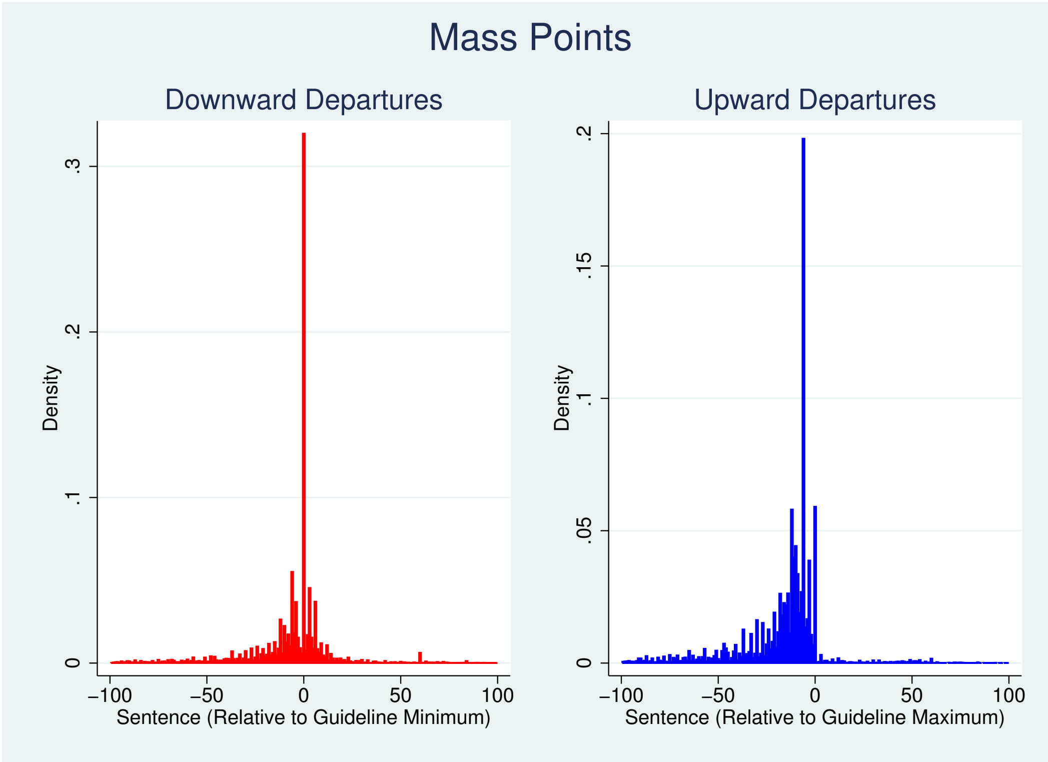
Table 5: Decomposition of Gender and Racial Sentencing Gaps

		Obs Gap	Pred Gap	Explained	Unexplained		
					Total	Male Pref	Female Pref
Male vs Female (Whites)	Guidelines	20.80	20.64	14.34	6.30	0.10	6.20
	No Guidelines		19.53	11.30	8.24	0.71	7.53
Male vs Female (Blacks)	Guidelines	48.73	46.07	37.53	8.54	2.68	5.86
	No Guidelines		43.22	31.33	11.89	3.60	8.29
Male vs Female (Hispancis)	Guidelines	24.81	23.84	17.41	6.43	0.49	5.94
	No Guidelines		22.74	14.61	8.14	-0.02	8.15
					Total	White Pref	Minority Pref
White vs Black (Males)	Guidelines	-32.03	-31.56	-28.98	-2.58	0.10	-2.68
	No Guidelines		-27.03	-24.14	-2.89	0.71	-3.60
White vs Hispanics (Males)	Guidelines	-9.05	-6.84	-6.45	-0.39	0.10	-0.49
	No Guidelines		-5.19	-5.91	0.72	0.71	0.02
White vs Black (Females)	Guidelines	-4.10	-6.13	-5.79	-0.33	-6.20	5.86
	No Guidelines		-3.34	-4.11	0.76	-7.53	8.29
White vs Hispanics (Females)	Guidelines	-5.04	-3.64	-3.38	-0.26	-6.20	5.94
	No Guidelines		-1.97	-2.60	0.63	-7.53	8.15

Table 6: Decomposition of Gender and Racial Sentencing Gaps (Base Offense Level)

		Obs Gap	Pred Gap	Explained	Unexplained		
					Total	Male Pref	Female Pref
Male vs Female (Whites)	Guidelines	20.80	20.03	12.13	7.89	-0.16	8.06
	No Guidelines		18.97	8.47	10.51	0.65	9.86
Male vs Female (Blacks)	Guidelines	48.73	45.73	33.45	12.28	4.80	7.48
	No Guidelines		41.05	24.50	16.55	6.03	10.52
Male vs Female (Hispancis)	Guidelines	24.81	23.18	14.91	8.27	0.17	8.10
	No Guidelines		21.36	10.88	10.47	-0.95	11.42
					Total	White Pref	Minority Pref
White vs Black (Males)	Guidelines	-32.03	-32.02	-27.05	-4.96	-0.16	-4.80
	No Guidelines		-25.39	-20.00	-5.39	0.65	-6.03
White vs Hispanics (Males)	Guidelines	-9.05	-7.47	-7.14	-0.34	-0.16	-0.17
	No Guidelines		-4.53	-6.12	1.59	0.65	0.95
White vs Black (Females)	Guidelines	-4.10	-6.32	-5.74	-0.57	-8.06	7.48
	No Guidelines		-3.31	-3.97	0.66	-9.86	10.52
White vs Hispanics (Females)	Guidelines	-5.04	-4.32	-4.36	0.04	-8.06	8.10
	No Guidelines		-2.15	-3.71	1.56	-9.86	11.42

Figure 1:


Mass Points
Downward Departures
Density
0
.1
.2
.3
–100
–50
0
50
100
Sentence (Relative to Guideline Minimum)
Upward Departures
Density
0
.05
.1
.15
.2
–100
–50
0
50
100
Sentence (Relative to Guideline Maximum)

A Supplementary Tables

Table A1: Summary Stats for White Males (33232 Observations)

variable	**mean**	**sd**	**min**	**max**
TOTALMON	36.03	51.48	0.00	783.00
HISCHOOL	0.46	0.50	0.00	1.00
GED	0.14	0.35	0.00	1.00
SOMECOLL	0.24	0.43	0.00	1.00
COLLGRAD	0.12	0.33	0.00	1.00
MARRD	0.25	0.43	0.00	1.00
NUMDEPEN	1.12	1.42	0.00	12.00
DEFENSEP	0.37	0.48	0.00	1.00
CRIMHIS2	0.12	0.33	0.00	1.00
CRIMHIS3	0.14	0.34	0.00	1.00
CRIMHIS4	0.07	0.26	0.00	1.00
CRIMHIS5	0.04	0.20	0.00	1.00
CRIMHIS6	0.10	0.29	0.00	1.00
XFOLSOR	17.06	8.13	0.00	50.00
XFOLSOR2	356.94	314.55	0.00	2500.00
XFOLSOR3	8566.77	10881.87	0.00	125000.00
AGE	37.91	11.24	18.00	65.00
AGE2	1563.09	898.74	324.00	4225.00
CIRC2	0.11	0.31	0.00	1.00
CIRC3	0.05	0.21	0.00	1.00
CIRC4	0.05	0.22	0.00	1.00
CIRC5	0.10	0.30	0.00	1.00
CIRC6	0.08	0.28	0.00	1.00
CIRC7	0.04	0.20	0.00	1.00
CIRC8	0.08	0.27	0.00	1.00
CIRC9	0.27	0.44	0.00	1.00
CIRC10	0.06	0.23	0.00	1.00
CIRC11	0.14	0.34	0.00	1.00
y1	0.13	0.33	0.00	1.00
y2	0.13	0.34	0.00	1.00
y3	0.13	0.34	0.00	1.00
y4	0.14	0.34	0.00	1.00
y5	0.14	0.35	0.00	1.00
y6	0.15	0.36	0.00	1.00

Table A2: Summary Stats for White Females (7488 Observations)

variable	**mean**	**sd**	**min**	**max**
TOTALMON	15.23	28.30	0.00	384.00
HISCHOOL	0.50	0.50	0.00	1.00
GED	0.12	0.32	0.00	1.00
SOMECOLL	0.28	0.45	0.00	1.00
COLLGRAD	0.06	0.25	0.00	1.00
MARRD	0.25	0.43	0.00	1.00
NUMDEPEN	1.10	1.30	0.00	10.00
DEFENSEP	0.30	0.46	0.00	1.00
CRIMHIS2	0.10	0.29	0.00	1.00
CRIMHIS3	0.10	0.30	0.00	1.00
CRIMHIS4	0.03	0.17	0.00	1.00
CRIMHIS5	0.02	0.13	0.00	1.00
CRIMHIS6	0.03	0.16	0.00	1.00
XFOLSOR	13.91	7.38	0.00	42.00
XFOLSOR2	248.09	260.58	0.00	1764.00
XFOLSOR3	5310.61	8380.19	0.00	74088.00
AGE	35.93	10.50	18.00	65.00
AGE2	1401.54	810.51	324.00	4225.00
CIRC2	0.08	0.27	0.00	1.00
CIRC3	0.04	0.19	0.00	1.00
CIRC4	0.05	0.22	0.00	1.00
CIRC5	0.12	0.33	0.00	1.00
CIRC6	0.09	0.29	0.00	1.00
CIRC7	0.04	0.20	0.00	1.00
CIRC8	0.09	0.29	0.00	1.00
CIRC9	0.28	0.45	0.00	1.00
CIRC10	0.06	0.24	0.00	1.00
CIRC11	0.13	0.34	0.00	1.00
y1	0.12	0.32	0.00	1.00
y2	0.13	0.34	0.00	1.00
y3	0.14	0.35	0.00	1.00
y4	0.15	0.36	0.00	1.00
y5	0.15	0.35	0.00	1.00
y6	0.15	0.36	0.00	1.00

Table A3: Summary Stats for Black Males (22834 Observations)

variable	**mean**	**sd**	**min**	**max**
TOTALMON	68.06	70.29	0.00	894.00
HISCHOOL	0.58	0.49	0.00	1.00
GED	0.13	0.34	0.00	1.00
SOMECOLL	0.19	0.39	0.00	1.00
COLLGRAD	0.03	0.18	0.00	1.00
MARRD	0.21	0.41	0.00	1.00
NUMDEPEN	1.57	1.77	0.00	22.00
DEFENSEP	0.20	0.40	0.00	1.00
CRIMHIS2	0.12	0.33	0.00	1.00
CRIMHIS3	0.19	0.39	0.00	1.00
CRIMHIS4	0.13	0.33	0.00	1.00
CRIMHIS5	0.08	0.27	0.00	1.00
CRIMHIS6	0.18	0.38	0.00	1.00
XFOLSOR	21.05	9.32	0.00	49.00
XFOLSOR2	529.96	394.30	0.00	2401.00
XFOLSOR3	14743.00	14597.41	0.00	117649.00
AGE	31.41	9.12	18.00	65.00
AGE2	1069.73	663.91	324.00	4225.00
CIRC2	0.10	0.30	0.00	1.00
CIRC3	0.06	0.23	0.00	1.00
CIRC4	0.12	0.33	0.00	1.00
CIRC5	0.11	0.32	0.00	1.00
CIRC6	0.13	0.34	0.00	1.00
CIRC7	0.05	0.21	0.00	1.00
CIRC8	0.07	0.26	0.00	1.00
CIRC9	0.12	0.32	0.00	1.00
CIRC10	0.03	0.16	0.00	1.00
CIRC11	0.20	0.40	0.00	1.00
y1	0.12	0.33	0.00	1.00
y2	0.13	0.34	0.00	1.00
y3	0.13	0.34	0.00	1.00
y4	0.14	0.35	0.00	1.00
y5	0.15	0.36	0.00	1.00
y6	0.16	0.36	0.00	1.00

Table A4: Summary Stats for Black Females (5167 Observations)

variable	mean	sd	min	max
TOTALMON	19.33	36.55	0.00	360.00
HISCHOOL	0.52	0.50	0.00	1.00
GED	0.07	0.26	0.00	1.00
SOMECOLL	0.31	0.46	0.00	1.00
COLLGRAD	0.05	0.22	0.00	1.00
MARRD	0.17	0.37	0.00	1.00
NUMDEPEN	1.56	1.54	0.00	13.00
DEFENSEP	0.17	0.38	0.00	1.00
CRIMHIS2	0.11	0.31	0.00	1.00
CRIMHIS3	0.11	0.31	0.00	1.00
CRIMHIS4	0.04	0.19	0.00	1.00
CRIMHIS5	0.02	0.14	0.00	1.00
CRIMHIS6	0.05	0.21	0.00	1.00
XFOLSOR	13.90	8.13	1.00	43.00
XFOLSOR2	259.11	296.73	1.00	1849.00
XFOLSOR3	5938.51	9839.43	1.00	79507.00
AGE	32.41	9.39	18.00	65.00
AGE2	1138.20	688.66	324.00	4225.00
CIRC2	0.10	0.30	0.00	1.00
CIRC3	0.05	0.22	0.00	1.00
CIRC4	0.10	0.30	0.00	1.00
CIRC5	0.14	0.34	0.00	1.00
CIRC6	0.16	0.36	0.00	1.00
CIRC7	0.04	0.20	0.00	1.00
CIRC8	0.06	0.24	0.00	1.00
CIRC9	0.13	0.34	0.00	1.00
CIRC10	0.02	0.15	0.00	1.00
CIRC11	0.18	0.39	0.00	1.00
y1	0.13	0.33	0.00	1.00
y2	0.13	0.33	0.00	1.00
y3	0.13	0.34	0.00	1.00
y4	0.15	0.36	0.00	1.00
y5	0.15	0.35	0.00	1.00
y6	0.16	0.37	0.00	1.00

Table A5: Summary Stats for Hispanic Males (13253 Observations)

variable	**mean**	**sd**	**min**	**max**
TOTALMON	45.08	53.80	0.00	600.00
HISCHOOL	0.67	0.47	0.00	1.00
GED	0.11	0.31	0.00	1.00
SOMECOLL	0.15	0.35	0.00	1.00
COLLGRAD	0.03	0.16	0.00	1.00
MARRD	0.32	0.47	0.00	1.00
NUMDEPEN	1.65	1.73	0.00	13.00
DEFENSEP	0.23	0.42	0.00	1.00
CRIMHIS2	0.12	0.33	0.00	1.00
CRIMHIS3	0.17	0.37	0.00	1.00
CRIMHIS4	0.08	0.27	0.00	1.00
CRIMHIS5	0.04	0.20	0.00	1.00
CRIMHIS6	0.07	0.25	0.00	1.00
XFOLSOR	19.46	8.21	1.00	46.00
XFOLSOR2	446.02	347.54	1.00	2116.00
XFOLSOR3	11454.01	12894.54	1.00	97336.00
AGE	31.51	9.97	18.00	65.00
AGE2	1092.41	730.05	324.00	4225.00
CIRC2	0.08	0.28	0.00	1.00
CIRC3	0.03	0.18	0.00	1.00
CIRC4	0.01	0.09	0.00	1.00
CIRC5	0.25	0.43	0.00	1.00
CIRC6	0.02	0.14	0.00	1.00
CIRC7	0.01	0.10	0.00	1.00
CIRC8	0.02	0.14	0.00	1.00
CIRC9	0.29	0.45	0.00	1.00
CIRC10	0.10	0.29	0.00	1.00
CIRC11	0.08	0.27	0.00	1.00
y1	0.09	0.29	0.00	1.00
y2	0.11	0.31	0.00	1.00
y3	0.11	0.31	0.00	1.00
y4	0.13	0.34	0.00	1.00
y5	0.16	0.37	0.00	1.00
y6	0.18	0.38	0.00	1.00

Table A6: Summary Stats for Hispanic Females (3017 Observations)

variable	**mean**	**sd**	**min**	**max**
TOTALMON	20.27	29.20	0.00	324.00
HISCHOOL	0.66	0.47	0.00	1.00
GED	0.10	0.30	0.00	1.00
SOMECOLL	0.17	0.38	0.00	1.00
COLLGRAD	0.02	0.16	0.00	1.00
MARRD	0.27	0.44	0.00	1.00
NUMDEPEN	1.74	1.56	0.00	12.00
DEFENSEP	0.19	0.40	0.00	1.00
CRIMHIS2	0.09	0.29	0.00	1.00
CRIMHIS3	0.08	0.27	0.00	1.00
CRIMHIS4	0.02	0.15	0.00	1.00
CRIMHIS5	0.01	0.10	0.00	1.00
CRIMHIS6	0.01	0.11	0.00	1.00
XFOLSOR	16.45	7.29	0.00	43.00
XFOLSOR2	323.64	274.36	0.00	1849.00
XFOLSOR3	7263.64	9139.77	0.00	79507.00
AGE	31.67	9.68	18.00	65.00
AGE2	1096.75	698.76	324.00	4225.00
CIRC2	0.08	0.27	0.00	1.00
CIRC3	0.03	0.18	0.00	1.00
CIRC4	0.01	0.09	0.00	1.00
CIRC5	0.26	0.44	0.00	1.00
CIRC6	0.02	0.15	0.00	1.00
CIRC7	0.01	0.09	0.00	1.00
CIRC8	0.02	0.13	0.00	1.00
CIRC9	0.32	0.47	0.00	1.00
CIRC10	0.11	0.31	0.00	1.00
CIRC11	0.07	0.25	0.00	1.00
y1	0.10	0.30	0.00	1.00
y2	0.09	0.29	0.00	1.00
y3	0.11	0.31	0.00	1.00
y4	0.14	0.35	0.00	1.00
y5	0.16	0.36	0.00	1.00
y6	0.19	0.39	0.00	1.00

Table A7: Expected Sentence in Months by Group and Regime

		With Own Weights		Pooled Weights	
		Y Hat	P Hat	Y Hat	P Hat
White M	Regime 1	0.00	0.24	0.00	0.25
	Regime 2	29.89	0.21	29.68	0.21
	Regime 3	43.63	0.31	43.63	0.31
	Regime 4	48.23	0.06	48.24	0.05
	Regime 5	57.30	0.10	57.30	0.09
	Regime 6	76.36	0.09	77.10	0.09
	Total	40.11		40.01	
White F	Regime 1	0.00	0.32	0.00	0.32
	Regime 2	17.03	0.23	17.93	0.13
	Regime 3	25.08	0.31	25.08	0.32
	Regime 4	28.28	0.04	28.68	0.04
	Regime 5	34.61	0.07	34.61	0.10
	Regime 6	44.91	0.04	56.12	0.09
	Total	19.47		25.66	
Black M	Regime 1	0.00	0.14	0.00	0.14
	Regime 2	53.98	0.27	55.12	0.30
	Regime 3	78.88	0.30	78.88	0.31
	Regime 4	85.80	0.07	85.36	0.07
	Regime 5	104.39	0.10	104.39	0.09
	Regime 6	122.25	0.11	116.60	0.09
	Total	71.66		68.98	
Black F	Regime 1	0.00	0.29	0.00	0.29
	Regime 2	21.51	0.20	22.19	0.14
	Regime 3	28.14	0.34	28.14	0.32
	Regime 4	32.00	0.04	32.41	0.05
	Regime 5	38.67	0.07	38.67	0.10
	Regime 6	51.67	0.05	61.05	0.10
	Total	25.60		31.46	
Hisp M	Regime 1	0.00	0.17	0.00	0.20
	Regime 2	36.52	0.26	35.63	0.25
	Regime 3	55.60	0.33	55.60	0.31
	Regime 4	59.70	0.06	59.87	0.06
	Regime 5	72.97	0.09	72.97	0.09
	Regime 6	86.28	0.08	88.88	0.08
	Total	46.95		46.46	
Hisp F	Regime 1	0.00	0.28	0.00	0.28
	Regime 2	19.88	0.29	20.67	0.18
	Regime 3	32.24	0.31	32.24	0.32
	Regime 4	35.61	0.04	36.10	0.05
	Regime 5	42.43	0.05	42.43	0.09
	Regime 6	52.33	0.04	63.86	0.08
	Total	23.11		29.05	

B Robustness Checks

Table B1: Decomposition of Gender and Racial Sentencing Gaps (Pre 1999)

		Obs Gap	Pred Gap	Explained	Unexplained		
					Total	Male Pref	Female Pref
Male vs Female (Whites)	Guidelines	20.02	20.82	14.53	6.29	0.20	6.09
	No Guidelines		18.97	10.94	8.03	0.64	7.38
Male vs Female (Blacks)	Guidelines	47.03	44.01	35.06	8.95	2.12	6.83
	No Guidelines		42.48	29.72	12.76	3.77	8.99
Male vs Female (Hispanics)	Guidelines	27.03	25.77	18.25	7.52	1.57	5.95
	No Guidelines		25.27	15.16	10.11	0.34	9.77
					Total	White Pref	Minority Pref
White vs Black (Males)	Guidelines	-32.34	-31.65	-29.74	-1.92	0.20	-2.12
	No Guidelines		-27.45	-24.32	-3.13	0.64	-3.77
White vs Hispanics (Males)	Guidelines	-13.67	-11.54	-10.18	-1.37	0.20	-1.57
	No Guidelines		-7.93	-8.24	0.30	0.64	-0.34
White vs Black (Females)	Guidelines	-5.33	-8.47	-9.20	0.74	-6.09	6.83
	No Guidelines		-3.94	-5.54	1.60	-7.38	8.99
White vs Hispanics (Females)	Guidelines	-6.66	-6.60	-6.46	-0.14	-6.09	5.95
	No Guidelines		-1.63	-4.02	2.39	-7.38	9.77

Table B2: Decomposition of Gender and Racial Sentencing Gaps (Post 1999)

		Obs Gap	Pred Gap	Explained	Unexplained		
					Total	Male Pref	Female Pref
Male vs Female (Whites)	Guidelines	20.88	20.04	14.03	6.01	-0.03	6.04
	No Guidelines		19.51	11.34	8.17	0.80	7.37
Male vs Female (Blacks)	Guidelines	50.96	48.11	39.56	8.54	3.18	5.36
	No Guidelines		44.30	33.03	11.27	3.28	7.99
Male vs Female (Hispanics)	Guidelines	23.70	22.87	17.21	5.66	-0.09	5.75
	No Guidelines		21.37	14.30	7.07	-0.29	7.36
					Total	White Pref	Minority Pref
White vs Black (Males)	Guidelines	-32.03	-31.37	-28.16	-3.21	-0.03	-3.18
	No Guidelines		-26.41	-23.93	-2.48	0.80	-3.28
White vs Hispanics (Males)	Guidelines	-5.95	-4.25	-4.31	0.06	-0.03	0.09
	No Guidelines		-3.16	-4.25	1.09	0.80	0.29
White vs Black (Females)	Guidelines	-1.94	-3.30	-2.62	-0.68	-6.04	5.36
	No Guidelines		-1.62	-2.24	0.62	-7.37	7.99
White vs Hispanics (Females)	Guidelines	-3.13	-1.43	-1.13	-0.29	-6.04	5.75
	No Guidelines		-1.30	-1.29	-0.01	-7.37	7.36

Table B3: Decomposition of Gender and Racial Sentencing Gaps (No Counsel Information)

		Obs Gap	Pred Gap	Explained	Unexplained		
					Total	Male Pref	Female Pref
Male vs Female (Whites)	Guidelines	20.04	20.06	13.58	6.47	-0.14	6.61
	No Guidelines		18.50	10.90	7.60	0.49	7.11
Male vs Female (Blacks)	Guidelines	53.76	49.86	40.60	9.26	2.84	6.42
	No Guidelines		46.79	34.43	12.36	3.61	8.75
Male vs Female (Hispancis)	Guidelines	27.65	26.07	19.48	6.58	0.37	6.22
	No Guidelines		23.92	16.47	7.45	-0.87	8.32
					Total	White Pref	Minority Pref
White vs Black (Males)	Guidelines	-37.32	-35.87	-32.89	-2.98	-0.14	-2.84
	No Guidelines		-30.52	-27.40	-3.12	0.49	-3.61
White vs Hispanics (Males)	Guidelines	-12.45	-9.64	-9.13	-0.51	-0.14	-0.37
	No Guidelines		-6.83	-8.19	1.36	0.49	0.87
White vs Black (Females)	Guidelines	-3.60	-6.07	-5.87	-0.20	-6.61	6.42
	No Guidelines		-2.23	-3.87	1.64	-7.11	8.75
White vs Hispanics (Females)	Guidelines	-4.84	-3.63	-3.23	-0.40	-6.61	6.22
	No Guidelines		-1.40	-2.62	1.21	-7.11	8.32

Table B4: Decomposition of Gender and Racial Sentencing Gaps (Circuit Time Trends)

		Obs Gap	Pred Gap	Explained	Unexplained		
					Total	Male Pref	Female Pref
Male vs Female (Whites)	Guidelines	20.80	20.64	14.36	6.28	0.09	6.19
	No Guidelines		19.53	11.32	8.21	0.69	7.52
Male vs Female (Blacks)	Guidelines	48.73	46.07	37.55	8.53	2.67	5.85
	No Guidelines		43.21	31.35	11.86	3.60	8.26
Male vs Female (Hispanics)	Guidelines	24.81	23.84	17.37	6.47	0.54	5.93
	No Guidelines		22.74	14.55	8.19	0.03	8.15
					Total	White Pref	Minority Pref
White vs Black (Males)	Guidelines	-32.03	-31.55	-28.97	-2.59	0.09	-2.67
	No Guidelines		-27.03	-24.12	-2.91	0.69	-3.60
White vs Hispanics (Males)	Guidelines	-9.05	-6.85	-6.40	-0.45	0.09	-0.54
	No Guidelines		-5.19	-5.84	0.66	0.69	-0.03
White vs Black (Females)	Guidelines	-4.10	-6.12	-5.78	-0.34	-6.19	5.85
	No Guidelines		-3.35	-4.08	0.74	-7.52	8.26
White vs Hispanics (Females)	Guidelines	-5.04	-3.65	-3.39	-0.26	-6.19	5.93
	No Guidelines		-1.98	-2.61	0.63	-7.52	8.15

Table B5: Decomposition of Gender and Racial Sentencing Gaps (Not including Co-operators)

		Obs Gap	Pred Gap	Explained	Unexplained		
					Total	Male Pref	Female Pref
Male vs Female (Whites)	Guidelines	23.35	23.02	19.32	3.70	0.46	3.24
	No Guidelines		22.00	15.93	6.07	1.38	4.68
Male vs Female (Blacks)	Guidelines	50.00	47.30	41.25	6.05	2.24	3.80
	No Guidelines		45.67	37.35	8.31	2.22	6.09
Male vs Female (Hispanics)	Guidelines	24.66	23.10	19.37	3.73	-1.09	4.82
	No Guidelines		23.03	17.70	5.34	-1.64	6.98
					Total	White Pref	Minority Pref
White vs Black (Males)	Guidelines	-30.35	-30.20	-28.42	-1.78	0.46	-2.24
	No Guidelines		-26.94	-26.10	-0.84	1.38	-2.22
White vs Hispanics (Males)	Guidelines	-7.61	-4.67	-6.22	1.55	0.46	1.09
	No Guidelines		-3.56	-6.59	3.02	1.38	1.64
White vs Black (Females)	Guidelines	-3.70	-5.93	-6.49	0.57	-3.24	3.80
	No Guidelines		-3.27	-4.68	1.41	-4.68	6.09
White vs Hispanics (Females)	Guidelines	-6.30	-4.59	-6.17	1.58	-3.24	4.82
	No Guidelines		-2.53	-4.82	2.29	-4.68	6.98

Table B6: Decomposition of Gender and Racial Sentencing Gaps (Including Trial Convictions)

		Obs Gap	Pred Gap	Explained	Unexplained		
					Total	Male Pref	Female Pref
Male vs Female (Whites)	Guidelines	22.98	22.37	15.60	6.77	0.09	6.69
	No Guidelines		22.17	12.70	9.47	0.81	8.66
Male vs Female (Blacks)	Guidelines	54.58	50.19	40.67	9.52	3.12	6.40
	No Guidelines		49.04	35.53	13.51	4.20	9.30
Male vs Female (Hispancis)	Guidelines	26.58	25.12	18.92	6.20	0.00	6.20
	No Guidelines		24.60	16.32	8.28	-0.89	9.17
					Total	White Pref	Minority Pref
White vs Black (Males)	Guidelines	-36.68	-35.18	-32.14	-3.03	0.09	-3.12
	No Guidelines		-31.24	-27.84	-3.40	0.81	-4.20
White vs Hispanics (Males)	Guidelines	-8.98	-6.75	-6.83	0.09	0.09	-0.00
	No Guidelines		-4.62	-6.32	1.70	0.81	0.89
White vs Black (Females)	Guidelines	-5.07	-7.36	-7.08	-0.28	-6.69	6.40
	No Guidelines		-4.36	-5.01	0.64	-8.66	9.30
White vs Hispanics (Females)	Guidelines	-5.37	-4.00	-3.51	-0.49	-6.69	6.20
	No Guidelines		-2.19	-2.70	0.51	-8.66	9.17

C Details on Econometric Model

The first set of boundary constraints on S_i arises from a downward departure from the guidelines:

$$U'(S_i \mid -\infty < \dot{S}_i < G_i^l) \leq 0$$
$$S_i \cdot U'(S_i \mid -\infty < \dot{S}_i < G_i^l) = 0$$

It follows that if the constrained utility maximizing value $\dot{S}_i \ \epsilon(-\infty, G_i^l)$, the actual sentence awarded is determined according to

$$\begin{aligned} S_i &= \max\left[0,\ \dot{S}_i = S_i^* + \theta_l\ \right] \\ &= \max\left[0,\ X_i'\beta + \theta_l + \varepsilon_i\right]. \end{aligned}$$

Thus, the empirical sentencing function is described by:

$$\begin{aligned} S_i &= X_i'\beta + \theta_l + \varepsilon_i \text{ if } 0 < RHS < G_i^l \\ &= 0 \text{ if } RHS \leq 0. \end{aligned}$$

The next set of boundary constraints occur in the interior region that encompasses non-departures from the guidelines.

$$U'(S_i \mid G_i^l < \dot{S}_i < G_i^h) = 0.$$

If the utility maximizing value $\dot{S}_i \ \epsilon(G_i^l,\ G_i^h\ )$, the empirical sentencing function is described by

$$\begin{aligned} S_i &= \dot{S}_i \\ &= S_i^* \\ &= X_i'\beta + \varepsilon_i \end{aligned}$$

Consider now the case for upward departures from the guidelines. If the utility maximizing value $\dot{S}_i > G_i^h$, it follows that

$$U'(S_i \mid G_i^h < \dot{S}_i\ < \infty) = 0.$$

In this case the empirical sentencing function is given by

$$\begin{aligned} S_i &= \dot{S}_i \\ &= S_i^* - \theta_h \\ &= X_i'\beta - \theta_h + \varepsilon_i. \end{aligned}$$

In order to accommodate mass points at G_i^l and G_i^h, we first need to determine the probabilities that the utility maximizing values $\dot{S}_i$ yield sentences that fall in the six regions already considered. From the assumption of a normal distribution on random utilities, it is easily shown that

$$\begin{aligned} prob(S_i = 0) &= prob(\varepsilon_i < -(X_i'\beta + \theta_l)) \\ &= 1 - \Phi\left(\frac{X_i'\beta + \theta_l}{\sigma_\varepsilon}\right) \end{aligned}$$

$$
\begin{aligned}
prob(0 < S_i < G_i^l) &= prob(S_i < G_i^l) - prob(S_i < 0) \\
&= \Phi\left(\frac{G_i^l - X_i'\beta - \theta_l}{\sigma_\varepsilon}\right) - \left[1 - \Phi\left(\frac{X_i'\beta + \theta_l}{\sigma_\varepsilon}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
prob(G_i^l < S_i \ < G_i^h) &= prob(S_i^* < G_i^h) - prob(S_i^* < G_i^l) \\
&= \Phi\left(\frac{G_i^h - X_i'\beta}{\sigma_\varepsilon}\right) - \Phi\left(\frac{G_i^l - X_i'\beta}{\sigma_\varepsilon}\right)
\end{aligned}
$$

$$
\begin{aligned}
prob(S_i > G_i^h) &= prob(\varepsilon_i > G_i^h - X_i'\beta + \theta_h) \\
&= 1 - \Phi\left(\frac{G_i^h - X_i'\beta + \theta_h}{\sigma_\varepsilon}\right).
\end{aligned}
$$

To determine the probability of a mass point at $S_i = G_i^l$, note

$$
\begin{aligned}
prob(S_i = G_i^l) &= prob(S_i^* < G_i^l) - prob(S_i < G_i^l) \\
&= \Phi\left(\frac{G_i^l - X_i'\beta}{\sigma_\varepsilon}\right) - \Phi\left(\frac{G_i^l - X_i'\beta - \theta_l}{\sigma_\varepsilon}\right).
\end{aligned}
$$

Similarly, the probability of a mass point at $S_i = G_i^h$ is determined according to

$$
\begin{aligned}
prob(S_i = G_i^h) &= prob(S_i^* > G_i^h) - prob(S_i > G_i^h) \\
&= \left[1 - prob(S_i^* < G_i^h\right] - \left[1 - prob(S_i < G_i^h)\right] \\
&= prob(S_i < G_i^h) - prob(S_i^* < G_i^h) \\
&= \Phi\left(\frac{G_i^h - X_i'\beta + \theta_h}{\sigma_\varepsilon}\right) - \Phi\left(\frac{G_i^h - X_i'\beta}{\sigma_\varepsilon}\right).
\end{aligned}
$$

It is readily verified that the probabilities over all regions sum to 1. We can summarize the six regions according to

Region 1:	$S_i = 0$
Region 2:	$0 < S_i < G_i^l$
Region 3:	$S_i = G_i^l$
Region 4:	$G_i^l < S_i < G_i^h$
Region 5:	$S_i = G_i^h$
Region 6:	$G_i^h < S_i$.

The corresponding log likelihood function for the sentencing model is specified by

$$
\begin{aligned}
ln\,(L) = & \sum_{i\epsilon\{S_i|S_i=0\}} ln\left[1-\Phi\left(\frac{X_i'\beta+\theta_l}{\sigma_\varepsilon}\right)\right] \qquad (3)\\
& + \sum_{i\epsilon\left\{S_i|S_i=G_i^l\right\}} ln\left[\Phi\left(\frac{G_i^l - X_i'\beta}{\sigma_\varepsilon}\right) - \Phi\left(\frac{G_i^l - X_i'\beta-\theta_l}{\sigma_\varepsilon}\right)\right]\\
& + \sum_{i\epsilon\left\{S_i|S_i=G_i^h\right\}} ln\left[\Phi\left(\frac{G_i^h - X_i'\beta+\theta_h}{\sigma_\varepsilon}\right) - \Phi\left(\frac{G_i^h - X_i'\beta}{\sigma_\varepsilon}\right)\right]\\
& + \sum_{i\epsilon\left\{S|_i 0<S_i<G_i^l\right\}} ln\left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{S_i - X_i'\beta-\theta_l}{\sigma_\varepsilon}\right)\right] + \sum_{G_i^l<S_i<G_i^h} ln\left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{S_i - X_i'\beta}{\sigma_\varepsilon}\right)\right]\\
& + \sum_{i\epsilon\left\{S_i|G_i^h<S_i\right\}} ln\left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{S_i - X_i'\beta+\theta_h}{\sigma_\varepsilon}\right)\right]
\end{aligned}
$$

where n = the number of observations for which $0 < S_i < G_i^l, G_i^l < S_i < G_i^h$, or $G_i^h < S_i$. We term this model a partially uncensored ordered probit model.

For each sentencing case there are six conditional sentences corresponding to each possible sentencing region:

$$E(S_i \mid S_i = 0) = 0$$

$$E(S_i \mid 0 < S_i < G_i^l) = X_i'\beta + \theta_l + \sigma_\varepsilon\left[\frac{\phi\left(\frac{-X_i'\beta-\theta_l}{\sigma_\varepsilon}\right) - \phi\left(\frac{G_i^l - X_i'\beta-\theta_l}{\sigma_\varepsilon}\right)}{\Phi\left(\frac{G_i^l - X_i'\beta-\theta_l}{\sigma_\varepsilon}\right) - \Phi\left(\frac{-X_i'\beta-\theta_l}{\sigma_\varepsilon}\right)}\right]$$

$$E(S_i \mid S_i = G_i^l) = G_i^l$$

$$E(S_i \mid G_i^l < S_i < G_i^h) = X_i'\beta + \sigma_\varepsilon\left[\frac{\phi\left(\frac{G_i^l - X_i'\beta}{\sigma_\varepsilon}\right) - \phi\left(\frac{G_i^h - X_i'\beta}{\sigma_\varepsilon}\right)}{\Phi\left(\frac{G_i^h - X_i'\beta}{\sigma_\varepsilon}\right) - \Phi\left(\frac{G_i^l - X_i'\beta}{\sigma_\varepsilon}\right)}\right]$$

$$E(S_i \mid S_i = G_i^h) = G_i^h$$

$$E(S_i \mid S_i > G_i^h) = X_i'\beta - \theta_h + \sigma_\varepsilon\left[\frac{\phi\left(\frac{G_i^h - X_i'\beta+\phi_h}{\sigma_\varepsilon}\right)}{1-\Phi\left(\frac{G_i^h - X_i'\beta+\phi_h}{\sigma_\varepsilon}\right)}\right].$$

The expected sentence for the ith case is calculated as

$$
\begin{aligned}
E(S_i) &= prob(S_i = 0) \cdot E(S_i \mid S_i = 0) + prob(0 < S_i < G_i^l) \cdot E(S_i \mid 0 < S_i < G_i^l) \\
&+prob(S_i = G_i^l) \cdot E(S_i \mid S_i = G_i^l) + prob(G_i^l < S_i < G_i^h) \cdot E(S_i \mid G_i^l < S_i < G_i^h) \\
&+prob(S_i = G_i^h) \cdot E(S_i \mid S_i = G_i^h) + prob(S_i > G_i^h) \cdot E(S_i \mid S_i > G_i^h) \\
&= prob(0 < S_i < G_i^l) \cdot E(S_i \mid 0 < S_i < G_i^l) + prob(S_i = G_i^l) \cdot \ G_i^l \\
&+prob(G_i^l < S_i < G_i^h) \cdot E(S_i \mid G_i^l < S_i < G_i^h) + prob(S_i = G_i^h) \cdot G_i^h \\
&+prob(S_i > G_i^h) \cdot E(S_i \mid S_i > G_i^h). \qquad (4)
\end{aligned}
$$

The estimated sentence for the ith individual ($\hat{S}_i$) is calculated by evaluating eq(4) at the estimated parameter values.