POLICY ESSAY

RACIAL DISPARITY IN WAKE OF THE *BOOKER/FANFAN* DECISION

Unwarranted disparity in the wake of the *Booker/Fanfan* decision

Implications for research and policy

Cassia Spohn
Arizona State University

In 1972, Marvin Frankel, U.S. District Judge for the Southern District of New York, issued an influential call for reform of the federal sentencing process (Frankel, 1972a, 1972b). Judge Frankel characterized the indeterminate sentencing system that existed at that time as "a bizarre 'nonsystem' of extravagant powers confided to variable and essentially unregulated judges, keepers, and parole officials" (Frankel, 1972b: 1; see also Gertner, 2007). Frankel was particularly concerned about the degree of discretion given to judges, which he maintained led to "lawlessness" in sentencing. Claiming that judges "were not trained at all" for "the solemn work of sentencing" (Frankel, 1972b: 6), Judge Frankel called for legislative reforms designed to regulate "the unchecked powers of the untutored judge" (Frankel, 1972b: 41). More to the point, he called for the creation of an administrative agency called a sentencing commission that would create rules for sentencing that judges would be required to follow.

Congress responded to Judge Frankel's call for reform of federal sentencing by enacting the Sentencing Reform Act of 1984 (SRA). The SRA created the United States Sentencing Commission (USSC), which was authorized to develop and implement presumptive sentencing guidelines designed to achieve "honesty," "uniformity," and "proportionality" in sentencing (USSC, 2001). The Act also abolished discretionary release on parole, stated that departures from the guidelines would be permitted only with written justification, and provided for appellate review of sentences to determine whether the guidelines were

Direct correspondence to Cassia Spohn, School of Criminology and Criminal Justice, Arizona State University, 411 N. Central Avenue, Phoenix, AZ 85004 (e-mail: Cassia.Spohn@asu.edu).

DOI:10.1111/j.1745-9133.2011.00768.x © 2011 American Society of Criminology

applied correctly or whether a departure was reasonable. The federal sentencing guidelines promulgated by the USSC went into effect in 1987.

The goals of those who championed the enactment of the federal sentencing guidelines varied. Liberals argued that structured sentencing practices would enhance fairness and equity in sentencing and would hold judges accountable for their decisions, whereas conservatives asserted that the reforms would lead to harsher penalties that eventually would deter criminal behavior. Reformers on both sides of the political spectrum, however, agreed that the changes were designed to curb discretion and reduce unwarranted disparity. Both conservatives and liberals urged sentencing reform as a means of reducing "lawlessness" (Frankel, 1972b) in sentencing. Reflecting this mindset, the *Federal Sentencing Guidelines Manual* states that one of the three objectives Congress sought to achieve in enacting the SRA was "reasonable uniformity in sentencing by narrowing the wide disparity in sentences imposed for similar criminal offenses committed by similar offenders" (USSC, 2001: Ch.1, Pt. A. 3). It is also clear that the guidelines were designed to eliminate discrimination based on legally irrelevant characteristics of the offender; in determining the appropriate sentence, judges are prohibited explicitly from considering the offender's race, sex, national origin, creed, religion, and socioeconomic status (USSC, 2001, Ch. 5, Pt. H1.10). As the Department of Justice stated in 1987, "unwarranted disparity caused by broad judicial discretion is the ill that the Sentencing Reform Act seeks to cure" (Trott, 1995: 197).

Although evidence on whether the federal sentencing guidelines were able to achieve their goal of eliminating unwarranted disparity in sentencing is mixed (for reviews, see Mitchell, 2005; Spohn, 2000; USSC, 2004), both social scientists (Miethe and Moore, 1985; Wooldredge, 2009) and legal scholars (Hofer, 2006) argued that the potential for racial and ethnic discrimination was limited by the fact that the guidelines were mandatory and that judge-initiated departures were regulated closely. These aspects of the guidelines, in other words, constrained judicial discretion and made it less likely that judges would resort to stereotypes linked to race/ethnicity and other legally irrelevant factors in attempting to fashion sentences that fit offenders and their crimes (for an alternative view, see Stith and Cabranes, 1998: 126–128).

A series of decisions by the U.S. Supreme Court that reshaped the state and federal sentencing process made the federal sentencing guidelines "effectively advisory" rather than mandatory (see *U.S. v. Booker,* 2005) and held that federal judges must make an individualized determination of the appropriate sentence based on the facts presented (see *Gall v. United States*, 2007) led several commentators (Frase, 2007; Hofer, 2007; Klein, 2005) to suggest that the result might be an increase in unwarranted disparity in federal sentencing. For example, Klein (2005: 720–721) noted that the *Booker* decision provided judges with the authority to fashion sentences that reflect the purposes of sentencing set forth in the SRA and argued that:

> [T]his authority, coupled with the admonition in 18 U.S.C. § 3661 that 'no limitations be placed on the information concerning the background, character, and conduct of a person convicted of an offense which a court of the United States may receive and consider for purposes of imposing an appropriate sentence,' allow trial judges *free reign* in gathering information and making discretionary sentencing decisions" (emphasis added).

Other scholars disagree, arguing that the Court's decision in *Booker* did not grant judges unfettered discretion at sentencing and that in the post-*Booker* era, the federal sentencing guidelines "remain as restrictive of judicial sentencing discretion as any system in the United States" (Reitz, 2005: 171; see also Berman, 2006; Bowman, 2005).

Advisory Guidelines and Unwarranted Disparity

Determining whether the Court's decisions have led to increased disparity in the federal sentencing process—and, more specifically, whether the decisions have increased disparity based on legally irrelevant offender characteristics such as race, ethnicity, and sex—clearly is important. In fact, Hofer (2007: 451) contended that this is "perhaps the most important empirical question facing policy makers." If, as critics of *Booker* and subsequent decisions assert, unwarranted disparity has increased as a result of the loosening of constraints on judicial discretion to determine the appropriate sentence and to depart from the guideline range, then some type of policy "fix" designed to refashion the guidelines as mandatory may be necessary. It also is a significant issue for sentencing scholars, whose work will inform the debate and influence subsequent policy changes. As Hofer (2007: 456) noted, our understanding of the impact of changes in sentencing policy must be "grounded in hard evidence and not in mere anecdotes or speculation."

Given the centrality of this issue, it is interesting that two major studies have reached different conclusions. Research by the USSC (2010) found that the offender's race, sex, citizenship status, education, and age affected the length of the prison sentence during at least some of the time periods under study. More to the point, that study's refined analysis found that disparity based on the *combination* of the offender's race and sex increased both post-*Booker* and post-*Gall*. To illustrate, compared with White male offenders, Black male offenders received 5.5% longer sentences in the pre-*Booker* period, but they received 15.2% longer sentences in the post-*Booker* period and 23.3% longer sentences in the post-*Gall* period. Hispanic male offenders received sentences that were 4.4% shorter than those imposed on White males in the pre-*Booke*r period, but they received 6.8% longer sentences than those imposed on White males in the post-*Gall* era. In contrast, the study by Ulmer, Light, and Kramer (2011, this issue), which (a) analyzed data over a longer period of time, (b) examined the likelihood of incarceration and the likelihood of departure as well as sentence length, (c) included somewhat different control variables, and (d) used a different

modeling strategy, found no increase in racial disparity in sentence length post-*Booker* and post-*Gall*. In fact, *less* disparity was found in sentence lengths for Black males (compared with White males) in the post-*Booker* and post-*Gall* period than found in 1994 and 1995, when the guidelines were mandatory and the standard for appellate review of sentences outside the guideline range was more stringent. Ulmer et al. also found, however, that though disparity in the odds of incarceration for Black male offenders relative to White male offenders was stable through the post-*Booker* period, it increased significantly post-*Gall.* Other important findings were that the increase in sentence length disparity for Black male offenders found by the USSC study was reduced substantially when a variable measuring the offender's criminal history points was included in the model, that a nontrivial portion of the sentence length disparity for Black males across all time periods was attributable to immigration offenses, and that greater disparity occurred in government-sponsored departures than in judge-initiated departures. Ulmer et al. concluded that their results suggest that Black male offenders' odds of imprisonment have increased post-*Gall* but call into question the "notion that *Booker* and *Gall* have caused increases in race/ethnic and gender sentence length disparity compared with the full range of years when the Guidelines were mandatory."

The contrasting findings of these two studies suggest that definitive answers to questions regarding the impact of the Supreme Court's *Booker* and *Gall* decisions remain elusive. They also suggest, consistent with previous work on the appropriate way to control for offense seriousness and criminal history in jurisdictions with presumptive guidelines (Engen and Gainey, 2000) and the use of a "total incarceration variable" combining prison and jail sentences (Holleran and Spohn, 2004), that the methodological choices made by researchers have important consequences for conclusions regarding the locus and extent of unwarranted disparity in sentencing. This finding, in turn, implies that if their work is to be relevant to policy and practice, then researchers whose focus is the federal sentencing process must agree on a general framework for analysis. It is not helpful to policy makers to be presented with the results of studies whose conflicting findings can be attributed simply to the fact that the authors used different methodologies and analytical strategies.

The results of the study conducted by Ulmer et al. (2011) have other implications for both research and policy. In terms of research, their study illustrates the importance of estimating models of the likelihood of incarceration and of the likelihood of a downward departure, as well as of the length of the prison sentence, as doing so can identify the locus of unwarranted disparity and point to appropriate policy remedies. If, as Ulmer et al. reveal, judges in the post-*Booker* era take the offender's race/ethnicity and sex into account primarily when deciding whether the offender should be sentenced to prison, it would make little sense to attempt to structure discretion by decreasing the range of sentences for each cell in the guideline matrix. Similarly, if unwarranted disparity is tied to government-initiated downward departures, especially those for providing substantial assistance, then an appropriate policy fix would be one targeting discretionary charging and plea bargaining

decisions by prosecutors. The sentence imposed on an offender who has been found guilty of a crime is the result of a collaborative exercise that involves criminal justice officials other than the judge. It results from a sentencing *process,* not from a single sentencing *decision*, and researchers must attempt to identify where in this process unwarranted disparity is found.

The study by Ulmer et al. (2011) also highlights the value of estimating models that include interaction effects and that test for indirect, as well as direct, effects. As numerous commentators have pointed out, simply attempting to determine whether race and ethnicity have direct effects on sentence outcomes is a theoretically unsophisticated and incomplete approach to a complex phenomenon. The more interesting question is, "When does the particular social characteristic matter—under what circumstances, for whom, and in interaction with what other factors?" (Wonders, 1996: 617). For example, a growing body of research suggests that findings of leniency for female offenders may be conditioned by race/ethnicity and that findings of preferential treatment for White offenders (or more punitive treatment for minorities) may be conditioned by sex (for a review of this research, see Brennan, 2009). Thus, a failure to consider the intersection of sex and race/ethnicity may result in inaccurate conclusions about the effects of these variables on sentencing outcomes. Both the USSC and Ulmer et al. test for the joint effects of the offender's race/ethnicity and sex, finding that Black and Hispanic male offenders are singled out for harsher treatment at various decision points. The same pattern—that is, harsher treatment of racial minorities—was not found for female offenders; in fact, Black female offenders faced lower odds of incarceration than White female offenders, and the lengths of the prison sentences imposed on Black, Hispanic, and White females were similar.

These results—coupled with Ulmer et al.'s (2011) findings that the offender's criminal history had a significant effect on sentence length, above and beyond its effect through the presumptive sentence, and that the effect was particularly pronounced (in all time periods) for Black male offenders—suggest that judges believe, first, that offenders with lengthier and more serious criminal histories deserve more punitive treatment; second, that the incorporation of the offender's criminal history score in the presumptive sentence does not account for this adequately; and, third, that Black male repeat offenders are particularly deserving of more punitive sentences. So, what are the policy implications of these findings? If the guidelines do not capture judges' assessments of the punitive impact of the offender's criminal history adequately, then perhaps the guideline matrix should be revised so that increases in the criminal history score result in larger increases in the range of presumptive sentences for each cell. The problem with this solution is that it would result in more punitive presumptive sentences for all offenders with more serious criminal histories, and we have no guarantee that judges would not continue to impose more severe sentences than the guidelines called for on repeat Black male offenders. If judges' assessments of offenders' dangerousness, threat, and likelihood of recidivism are tied to their evaluations of offenders' criminal histories, and if this connection is more salient for Black male offenders than for

other offenders, then tinkering with the guideline matrix in this way will not reduce the racial disparity in sentence outcomes.

Ulmer et al.'s (2011) study also illustrates the importance of disaggregating the data by offense type. They point out correctly that immigration cases, which now represent more than 25% of all cases adjudicated in federal courts, present unique challenges, are concentrated in a handful of districts, and are handled in distinct ways from most other offenses. Their results revealed that inclusion of these offenses resulted in larger estimates of racial/ethnic and gender disparity in sentence length than would be the case if these offenses were excluded, particularly for Black and Hispanic male offenders. This finding is important in that it provides policy makers with information regarding the circumstances under which, and the contexts in which, legally irrelevant offender characteristics come into play. It also suggests that researchers interested in pinpointing the location of unwarranted disparity in the federal sentencing process should analyze immigration cases, drug trafficking cases, fraud cases, and firearms cases separately. As Hawkins (1986–1987: 724) pointed out more than two decades ago, it is overly simplistic to assert that Blacks and Hispanics will receive "more severe punishment than whites for all crimes, under all conditions, and at similar levels of disproportion over time."

The Future of Federal Sentencing

Shortly after the Supreme Court ruled in the *Booker* case that the federal sentencing guidelines would be advisory rather than mandatory, calls came for Congress to act to "mandatorize" the guidelines (see House Bill H.R. 1528, which would have made the federal sentencing guidelines a series of mandatory minimum penalties). Other commentators, including the American Bar Association (2005), argued that any type of "*Booker*-fix" was premature and urged Congress to allow the advisory guidelines to operate for at least 1 year before enacting legislation designed to constrain judicial discretion at sentencing.

The advisory guideline system put into place by *Booker* (and reiterated by *Gall*) has now been in operation for more than 6 years. The result has not been the instability and lack of uniformity that critics of *Booker* predicted. In fact, a report by the USSC (2006) on the first post-*Booker* year of sentencing revealed that judges continued to consult and apply the guidelines, with the result that a within-guideline sentence was imposed in nearly two out of every three cases. Moreover, the severity of sentences imposed did not change substantially. According to Berman (2006: 157), the reality is that *Booker* "has not radically altered many central features of the federal sentencing system."

In 2006, a series of articles designed to "take stock concerning the state and direction of federal sentencing" (Berman, 2006: 157) appeared in the *Federal Sentencing Reporter*. The authors of one of these articles, both of whom were district court judges, commented on the remarkable stability in sentencing in the post-*Booker* period and noted that it

was "impossible to determine" whether the variations from the guidelines that did exist represented unwarranted disparity (Adelman and Deitrich, 2006: 160). The authors concluded that "at the present time, the argument that mandates are needed to eliminate unwarranted disparity is utterly unsupported by evidence" (2006: 160).

The question, of course, is whether this conclusion is still valid 5 years later. As the study conducted by Ulmer et al. (2011) makes clear, the answer to this question depends to some extent on the methodological and analytical choices made by researchers attempting to tease out the effects of offender race, ethnicity, and sex on sentence outcomes. One study (USSC, 2010) concluded that the racial/ethnic disparity in sentence length increased substantially in the post-*Booker* (for Black males) and the post-*Gall* periods (for both Black males and Hispanic males), but another study by researchers who made different methodological choices and whose findings challenge the conclusion of increasing unwarranted disparity in the federal sentencing process (Ulmer et al.) provides a different conclusion. However, it is important to point out that the authors of this latter study *do not* conclude that no unwarranted disparity exists in the federal sentencing process or that disparities have not surfaced for some types of offenders (i.e., Black male offenders, especially those with more serious criminal histories), some types of decisions (i.e., the decision to incarcerate or not), and some types of offenses (i.e., immigration offenses). Although it may be true that *Booker* and *Gall* have not "caused increases in race/ethnic and gender sentence length disparity compared with the full range of years when the Guidelines were mandatory" (Ulmer et al.), it does not necessarily follow from this that the now-advisory guidelines are being implemented without any consideration of legally irrelevant offender characteristics. As Ulmer et al. point out, the situations in which racial disparity surfaced in their study "warrant additional scrutiny and perhaps discussions of policy changes targeted specifically to those two circumstances."

References

American Bar Association, Criminal Justice Section. 2005. Report and recommendation on *Booker*. Reprinted in *Federal Sentencing Reporter*, 17: 334–337.

Adelman, Lynn and Jon Deitrich. 2006. Disparity: Not a reason to "fix" *Booker*. *Federal Sentencing Reporter*, 18: 160–163.

Berman, Douglas. 2005. Forward: Beyond *Blakely* and *Booker*: Pondering modern sentencing process. *Journal of Criminal Law and Criminology*, 95: 653–688.

Berman, Douglas. 2006. Now what? The post-*Booker* challenge for Congress and the courts. *Federal Sentencing Reporter*, 18: 157–159.

Bowman, Frank O. 2005. Beyond band aids: A proposal for reconfiguring federal sentencing after *Booker*. *University of Chicago Legal Forum*, 2005: 149–216.

Brennan, Pauline K. 2009. The joint effects of offender race/ethnicity and sex on sentencing outcomes. In (Marvin D. Krohn, Alan J. Lizotte, and Gina Penly Hall, eds.), *Handbook of Crime and Deviance*. New York: Springer Science and Business Media.

Engen, Rodney L. and Randy R. Gainey. 2000. Modeling the effects of legally relevant and extralegal factors under sentencing guidelines: The rules have changed. *Criminology*, 38: 1207–1229.

Frankel, Marvin E. 1972a. *Criminal sentences: Law without order*. New York: Hall and Wang.

Frankel, Marvin E. 1972b. Lawlessness in sentencing. *University of Cincinnati Law Review*, 41: 1–54.

Frase, Richard S. 2007. The *Apprendi-Blakely* cases: Sentencing reform counter-revolution? *Criminology & Public Policy*, 6: 403–432.

Gertner, Nancy. 2007. From omnipotence to impotence: American judges and sentencing. *Ohio State Journal of Criminal Law*, 4: 523–529.

Hawkins, Darnell F.. 1986–1987. Beyond anomalies: Rethinking the conflict perspective on race and criminal punishment. *Social Forces*, 65: 719–745.

Hofer, Paul J. 2006. Immediate and long-term effects of *United States v. Booker*: More discretion, more disparity, or better reasoned sentences? *Arizona State Law Journal*, 38: 425–468.

Holleran, David and Cassia Spohn. 2004. On the use of the total incarceration variable in sentencing research. *Criminology* 42: 211–240.

Hofer, Paul J. 2007. *United States v. Booker* as a natural experiment: Using the empirical research to inform the federal sentencing policy debate. *Criminology & Public Policy*, 6: 433–460.

Klein, Susan R. 2005. The return of federal judicial discretion in criminal sentencing. *Valparaiso University Law Review*, 39: 693–740.

Miethe, Terance D. and Charles Moore. 1985. Racial differences in criminal processing: The consequences of model selection on conclusions about differential treatment. *Journal of Criminal Law and Criminology*, 78: 155–176.

Mitchell, Ojmarrh. 2005. A meta-analysis of race and sentencing research: Explaining the inconsistencies. *Journal of Quantitative Criminology*, 21: 439–466.

Reitz, Kevin R. 2005. The enforceability of sentencing guidelines. *Stanford Law Review* 58: 155–174.

Spohn, Cassia. 2000. Thirty years of sentencing reform: The quest for a racially neutral sentencing process. In *Policies, Processes and Decisions of the Criminal Justice System*, Vol. 3, *Criminal Justice 2000*. Washington, DC: U.S. Department of Justice.

Stith, Kate and Jose A. Cabranes. 1998. *Fear of Judging: Sentencing Guidelines in the Federal Courts*. Chicago: University of Chicago Press.

Trott, Stephen S. 1995. Letter to the Honorable William W. Wilkins, Jr. Reprinted in the *Federal Sentencing Reporter*, 8: 196–198.

Ulmer, Jeffery T, Michael T. Light, and John H. Kramer. 2011. Racial disparity in the wake of the *Booker/Fanfan* decision: An alternative analysis to the USSC's 2010 report. *Criminology & Public Policy*. This issue.

United States Sentencing Commission. 2001. *Federal Sentencing Guidelines Manual*. St. Paul, MN: West.

United States Sentencing Commission. 2004. *Fifteen Years of Guidelines Sentencing: An Assessment of How Well the Federal Criminal Justice System Is Achieving the Goals of Sentencing Reform*. Washington, DC: Author.

United States Sentencing Commission. 2006. *Report on the Impact of United States v. Booker on Federal Sentencing*. Washington, DC: Author.

United States Sentencing Commission. 2010. *Demographic Differences in Federal Sentencing Practices: An Update of the Booker Report's Multivariate Regression Analysis*. Washington, DC: Author.

Wonders, Nancy A. 1996. Determinate sentencing: A feminist and postmodern story. *Justice Quarterly*, 13: 611–648.

Wooldredge, John. 2009. Short- versus long-term effects of Ohio's switch to more structured sentencing on extralegal disparities in prison sentences in an urban court. *Criminology & Public Policy*, 8: 285–312.

Court Cases Cited

U.S. v. Booker, 543 U.S. 220 (2005).

Gall v. United States, 128 S. Ct. 586 (2007).

Statutes Cited

Consumer Privacy Protection Act of 2005, H.R. 1528.

Sentencing Reform Act, P.L. No. 98–473, 98 Stat. 1987 (1984).

Cassia Spohn is a Professor in the School of Criminology and Criminal Justice at Arizona State University, where she also serves as the director of the doctoral program. She is the coauthor of five books, including *The Color of Justice: Race, Ethnicity, and Crime in America* and *How Do Judges Decide? The Search for Fairness and Justice in Punishment.* Dr. Spohn has published extensively on prosecutors' charging decisions in sexual assault cases, the effect of race/ethnicity and gender on sentencing decisions, sentencing of drug offenders, and the deterrent effect of imprisonment. She is currently working on an NIJ-funded study of police decision making in sexual assault cases in Los Angeles.