New Federal Sentencing Data: Comparing Chalk and Cheese

For many practitioners who experienced the long period when the mandatory U.S. Sentencing Guidelines exerted near total dominion over the field of federal sentencing, and who greeted with cautious optimism the guidelines’ post-Booker demotion to their current advisory role, there is an abiding concern that this era will not last. A wary eye is trained on developments in the world of federal sentencing, with nagging worries about what they portend for the future. In this context, this article discusses the recent public release of federal sentencing data on a judge-specific basis, and a related report, issued March 5, 2012, bearing the headline “Surprising Judge-to-Judge Variations Documented in Federal Sentencing,” which has received significant media attention.1

One of the primary motivations for the promulgation of the guidelines pursuant to the Sentencing Reform Act of 1984 was the goal of eliminating unwarranted sentencing disparities among similar defendants found guilty of similar criminal conduct. The concept is much more easily stated than applied. Determining which defendants are “similar” and what is “similar” criminal conduct is at the heart of the sentencing process.

A primary lesson of the dissatisfaction with the mandatory guidelines era is that a sentencing system will not ultimately be seen as just if it has the effect of disregarding any of the untold number of variables that can, in a given case, be found persuasive in deeming two defendants or their conduct “dissimilar.” Determining which variables to focus upon and how much weight to give them in a particular case is the essence of judging when it comes to sentencing. Because of the complexity and subjectivity of the endeavor, no two judges will see every case the same way. As Judge John Gleeson observed in his thoughtful opinion in United States v. Ovid, “[t]here is nothing surprising or disturbing about the fact that once judgment is allowed to play a role in sentencing, it will be exercised differently by different people.”2

It is one thing to acknowledge that a properly functioning system will inevitably entail some sentencing variation among individual judges. But the notion that the identity of the judge is the main driver of vast disparities in the sentences of similarly situated defendants—“the claim that judges are inexplicably and unjustifiably all over the lot, sentencing based on their personal preferences”—is quite another.3 That undoubtedly would be seen as the kind of “unwarranted” disparity that cries out for the revival of some sort of mandatory system that limits judicial discretion.4

Experienced participants in the criminal justice system have been quick to point out the report’s manifest analytical flaws.

“The need to avoid unwarranted sentence disparities” among similarly situated defendants is one of the factors judges are required to consider under current law, pursuant to 18 U.S.C. §3553(a) (6). But the U.S. Court of Appeals for the Second Circuit has held that beyond considering the guidelines, this section imposes on sentencing judges only the “modest” additional duty of being “mindful of the general goal, however elusive, of national consistency.”5 Thus, as a practical matter, when a defendant seeks to challenge his sentence based on disparities with co-defendants, so long as there is a logical difference among them that is consistent with the broad sentencing factors stated in 18 U.S.C. § 3553(a), the sentence will be upheld.6

ROBERT J. ANELLO and RICHARD F. ALBERT are partners at Morvillo, Abramowitz, Grand, Iason, Anello & Bohrer. GRETCHAN R. OHLIG, an attorney, assisted in the preparation of this article.

Judge-Specific Data

Now into the debate comes the first ever public release of comprehensive judge-specific federal sentencing data. The data was obtained by the Transactional Records Access Clearinghouse (TRAC), an organization based at Syracuse University that gathers data on the federal government. For years, the Administrative Office of the U.S. Courts has gathered data on sentences imposed across the federal court system and provided it to the U.S. Sentencing Commission. Until now, categorization of this data on a judge-by-judge basis has not been publicly available, even though the commission has been understood to have it and to have made it available to individual judges upon request. TRAC obtained the data under the Freedom of Information Act after a 15-year legal battle with the government.

TRAC was founded in 1989 by co-directors David Burnham, a former investigative writer and New York Times reporter, and Susan Long, a statistician and professor in Syracuse University’s School of Management. TRAC’s declared purpose is “to provide the American people—and institutions of oversight like the Congress, public interest groups, and news organizations—with comprehensive and authoritative information about the operations of the federal government.”7

TRAC has developed an interactive database tool that makes available fairly detailed judge-specific sentencing information for more than 370,000 federal criminal cases completed during the period of fiscal years 2007 to 2011, including comparisons against other judges in the same district, and comparisons of each district against nationwide average and median sentences. Data for the sentences imposed by each of the 885 district judges who sentenced at least 50 defendants during the five-year period are available, segregated by broad category of cases, such as “drugs” or “white collar.” The tool also allows the database user to drill down into the underlying data, providing detailed information on individual cases. This information includes many of the same types of information typically found in a final judgment or docket sheet, such as the lead investigative agency, the name of the lead prosecutor, the identity of each count charged by statute, the dis-
position of each count, and the dates of key events in the case.

The database tool could conceivably be useful to practitioners seeking to identify cases for potential comparison, but because of the way the data is organized and what is omitted, it is difficult to see how it matters more than a possible starting point. The data seems potentially more useful to scholars, but scholars are not necessarily the primary target audience. The database tool is packaged to sell to law firms willing to pay an annual subscription fee, which escalates based on firm size.

Analysis Claims Variations

If one were trying to generate buzz to help sell paid subscriptions to the new database tool, one option would be to use the tool to generate a provocative analytical report, and then to issue a press release. And it appears that TRAC has succeeded in that, at least, through its March 5, 2012, report. The report states that an analysis conducted by TRAC using its new tool revealed “extensive and hard-to-explain differences in the sentencing practices by the judges working in many federal districts” and “indicates that the typical sentence handed down by a federal district court judge can be very different than the typical sentences handed down for similar cases by other judges in that same district.”

In analyzing so-called “white collar” cases, the report ranks individual federal districts according to the sentencing disparity among judges within those districts. The Northern District of Illinois tops the list, with a difference of 39 months in the median sentences imposed by the eight judges in the district. The Southern District of New York is the eighth highest on the list, with a difference of 22.5 months between the lowest median sentence for an applicable judge (0 months) and the highest (22.5 months); the Eastern District of New York ranks 17th on the list with a difference in median sentences of 12.2 months. The report finds that although differences in median white-collar sentences were within six months in approximately one-third of the 64 federal districts, in over a quarter of the districts, the difference between the median sentence assigned by each judge varied by over 18 months, and in 5 percent of the districts, the difference was over 24 months.

Media accounts have seized upon the TRAC report’s claims of a “wide sentencing disparity.” But experienced participants in the criminal justice system have been quick to point out the report’s manifest analytical flaws. A Fact Sheet released by the Sentencing Resource Council of the Office of Defender Services states that although “[t]he data released by TRAC might in the future shed light on federal sentencing...its initial analyses, and media coverage, demonstrate the danger of a little knowledge about a complex subject. TRAC’s analysis fails to meet minimal academic standards and should not be a basis for policy making.”

The central analytical weakness is the one that matters most—the failure to make any meaningful effort to ensure that the comparison is among cases that are, in fact, “similar” in the ways that actually matter to sentencing. The report groups sentences in broad, gross categories of offenses that are far too ill-defined and inclusive to be meaningful. The “white collar” category would lump together massive public company accounting fraud cases and Ponzi schemes with food stamp fraud, theft by a bank teller, and prosecutions arising from a defendant’s retention of Social Security payments intended for a deceased family member.

Further, TRAC’s analysis does not consider even the most rudimentary distinctions among offenses or defendants in these overall categories—the kind of broad distinctions the guidelines are properly criticized for relying upon too heavily, such as the loss amount in a fraud case or the defendant’s general role in the offense or whether he has an extensive prior criminal record. The analysis also does not make any distinctions based on the nature of any plea agreement having an impact on sentencing, including whether the government made a downward departure motion because of the defendant’s cooperation.

The report’s approach also wrongly assumes that over a five-year period, judges in the same district are likely to have a similar mix of cases. It disregards the massive distortions to the data that would be expected if a single judge were assigned a large, multi-defendant case or series of cases arising from a single investigation, where all or many of the cases could have similar features—such as weak evidence, many cooperators, or special plea agreement considerations, to name a few—that would properly skew multiple sentences significantly lower or higher than otherwise. Judge Nancy Gertner, a former federal judge in Boston and authority on federal sentencing, has aptly observed that the data upon which the TRAC report relies “doesn’t begin to address caseload or prosecutorial decisions. It doesn’t begin to address disparity.”

Conclusion

Near the outset of TRAC’s report, under the heading “Caution,” is the statement that, “A key requirement for achieving justice is that the judges in a court system have sufficient discretion to consider the totality of the circumstances in deciding that a sentence in a specific case is ‘just.’” TRAC claims that its goal is not “a lockstep sentencing system,” but to provide “both the courts and the public with accurate information” so that they can examine whether the goal of equal justice under the law is being achieved.

TRAC may well have taken a step toward its stated goals by making judge-specific sentencing data accessible, thereby aiding future scholarly study. It remains to be seen how valuable the database tool developed by TRAC will prove to be. But particularly in today’s world, no government institution can sustain itself by shielding its data from public scrutiny. Indeed, one federal judge has suggested that since “the cat is out of the bag,” perhaps it is time for the Sentencing Commission to reveal more detailed analyses to ensure judge-specific sentencing data is used properly.

But the initial analysis contained in TRAC’s report does not advance TRAC’s stated goals. The fundamental analytical flaws and overstated conclusions in the report, and the publicity generated by it, illustrate the need for vigilance to ensure that judge-specific sentencing data and other data like it are analyzed critically and intelligently. And if there is hope to avoid the policy mistakes of the past, it behooves those with real experience in how the federal criminal justice system works on a day-to-day basis to make sure they are heard.

---

4. United States v. Wilts, 476 F.3d 103, 110 (2d Cir. 2007), abrogated on other grounds, United States v. Cavena, 550 F.3d 180 (2d Cir. 2008) (en banc).
5. Id.; see, e.g., United States v. Exner, 431 Fed. Appx. 23 (2d Cir. 2011) (co-defendants not similar because one provided substantial assistance to the government); United States v. Memendez, 600 F.3d 253 (2d Cir. 2010) (disparities not unwarranted where co-defendants either pled guilty or were exceptionally honest in admitting to the crime and brought unique facts to court’s attention).
9. TRAC Report at Table 1.
11. Id. at p. 2.