Send to: Gertner, Nancy
HARVARD UNIVERSITY LAW SCHOOL
1511 MASSACHUSETTS AVE
CAMBRIDGE, MA 02138-2996

Time of Request: Wednesday, February 27, 2013 13:51:29 EST
Client ID/Project Name:
Number of Lines: 268
Job Number: 2828:396734949

Research Information

Service: EasySearch(TM) Feature
Print Request: Current Document: 19
Source: US Law Reviews and Journals, Combined
Search Terms: Federal Sentencing Reporter, Gertner

Federal Sentencing Reporter

TRAC DATA ON FEDERAL SENTENCING: Judge Identifiers, TRAC, and a Perfect World

25 Fed. Sent. R. 46

JUDGE NANCY GERTNER (RET.), Professor, Harvard Law School

October, 2012

LENGTH: 4588 words

TEXT:

In a perfect world, the recent release of the sentencing statistics of individual judges by the Transactional Records Access Clearinghouse of Syracuse University (TRAC) would not have raised an eyebrow (notwithstanding the provocative headline of the New York Times, "Wide Sentencing Disparity Found Among U.S. Judges). n1 A rational discourse about the factors influencing sentencing outcomes would have followed: How much of the apparent disparity in sentencing outcomes is attributable to the fact that different judges interpret Gall v. United States n2 and its progeny differently? Are there judges who simply sentence within the Guidelines no matter what, as if the seismic shifts in sentencing occasioned by United States v. Booker n3 had not happened? To what degree do appellate court decisions contribute to these patterns? Do the outcomes reflect the individual predilections of judges or prosecutors? Does the prosecutor charge differently in front of different judges (more mandatory minimum charges in front of judges who are perceived to be more lenient, for example)? What are the policies of different prosecutors with respect to cooperation agreements? How consistent are prosecutorial decisions within the same office?

n1 Mosi Secret, Wide Sentencing Disparity Found Among U.S. Judges, N.Y. Times, Mar. 5, 2012. TRAC announced the data in a press release trumpeting "Wide Variations Seen in Federal Sentencing," available at http://trac.syr.edu/whatsnew/email.120305.html.
n2 552 U.S. 38 (2007).
n3 543 U.S. 220 (2005).

And once sentencing factors causing apparently disparate outcomes were identified, we would debate whether they were lawful and appropriate. Were these differences caused by fair disputes about sentencing law or less appropriate biases of the judges? Are these warranted disparities, after all, sentences imposed by judges who are titrating their sentencing outcomes to real differences between offenders, or are they unwarranted, based on unexamined and less than transparent subjective factors, even wholly illegitimate factors, like race?

In fact, in the perfect world, those statistics would measure not only sentencing outcomes--the sentence that the defendant received, as if that is the end of the discussion--but also sentencing results: What program worked to reduce recidivism? What combination of punishments assists a long-term addict in maintaining sobriety? What happens to the numbers of young men sentenced to lengthy terms for crack-cocaine distribution after their release? How should that affect the sentence in the first instance?

In my perfect world, the statistics would not only help scholars and the public understand the nuances of federal sentencing, they would be invaluable to judges in making sentencing decisions themselves. They would form a part of sentencing information systems, which many scholars and judges (including me) have strongly recommended. n4 Judges would look carefully at the statistics. They would be eager to go over the data with a fine-tooth comb, discuss why their sentences were alike or different, engage with their colleagues and perhaps persuade them of the rightness of their approach, or the opposite, change their minds. They would analyze what worked and what did not work. Their sentences would be more consistent because they had the data to enable them to situate the case before them in the context of larger sentencing patterns in the district or in the country. They would behave, in short, like some physicians who study medical outcomes in a systematic way, or who, at the very least, compare therapies with their colleagues in peer review procedures. n5

n4 Marc L. Miller, A Map of Sentencing and a Compass for Judges: Sentencing Information Systems, Transparency, and the Next Generation of Reform, 105 Colum. L. Rev. 1351 (2005); James E. Felman, The State of the Sentencing Union: A Call for Fundamental Reexamination, 20 Fed. Sent'g Rep. 337, 338 (2008); J.C. Oleson, Blowing Out the Candles: A Few Thoughts on the 25th Anniversary of the Sentencing Reform Act, 45 U. Rich. L. Rev. 693, 744 (2011). In articles and in testimony before the Commission, I urged it to be more than simply the guideline police, but also to provide judges with the data with which they can exercise their new post-Booker authority. See Judge Nancy Gertner, Supporting Advisory Guidelines, 3 Harv. L. & Pol'y Rev. 261, 279-80 (2009).

n5 As suggested by Oleson, supra note 4, at 749-50, "in consigning people to prison, judges impose their sentences based on less information about what works and what is cost-effective than physicians who prescribe medicine. . . . Not only should [judges] be able to see what other judges had done when confronted with an analogous case--allowing a kind of common law of sentencing to flourish--but they could know whether a given sentence had worked."

In a perfect world, the statistics would be sophisticated, carefully measuring the myriad factors that judges were taking into account in their sentences now that the federal Guidelines are advisory. They would surely not be the simplistic reports found on the U.S. Sentencing Commission web page concerning departure or variance rates from the flawed Sentencing Guidelines, or which district's average sentences were high or low. n6 They would identify and try to quantify (to the extent possible) the rationale for the sentences--rationale that, in a post-Booker world, necessarily extends beyond the Guideline categories. They would identify differences between offenses, even when those differences do not fit the Guideline grid, for example, the difference between the crack offender who deals a given quantity of drugs out of the car in which he lives, and the one who deals the same quantity to pay for a new car. They would note differences between offenders as well--education, age, family background, addiction, mental health--even if those differences were once factors that the Sentencing Guidelines prohibited a judge from considering. n7

n6 The Sentencing Commission also provides more complex reports than the departure/variance or average sentence report, but their data is limited by the reports they receive, as described above.

n7 Compare the data available on the United States Sentencing Commission web site, www.ussc.gov with that available in the Victoria, Australia, Sentencing Advisory Council, http://www.sentencingcouncil.vic.gov.au.

But, alas, this is not a perfect world--not even close. The statistics, on which the TRAC data is based, like those of the Sentencing Commission, are flawed. Flawed statistics not only skew the public and even the scholarly discussion of sentencing, they distort judicial behavior. If the statistics actually measured the variety of factors on which sentencing was based, not only averages, medians, or departure rates, they would encourage judges to provide fuller explanations for their sentences--why this case was different from the case before it or the case in the next courtroom, what factors were really salient. But if the statistics are rough, judges are identified only in terms of their average sentences or worse, their departure/variance rates, some judges may well decide to stick to the Guidelines, lest they be unfairly criticized--as they surely will.

TRAC got access to raw sentencing data with judge identifiers, sentences imposed by 885 judges in more than 370,000 cases decided from 2007 through 2011. They provided the name of each federal judge, the number of defendants sentenced, the processing time, and the median and average sentences imposed for given categories. But among other problems, they did not account for the "why" of sentencing, the reasons for judges' decisions, the factors considered, the criterion relied on, and thus whether the disparities noted were warranted by meaningful distinctions between offenders. n8 Indeed, they could not. Except for one district--my own, Massachusetts--they did not have had access to the document in which the judge supposedly outlined the factors on which the sentence was based, the form entitled "Statement of Reasons" (SOR). I say "supposedly" because--sadly--even if they had gotten access to the form, or focused on the Massachusetts data, they could not have done much with it.

n8 For example, according to the Federal Public Defender Fact Sheet, the cases were categorized in broad categories like "drug" or "white collar." First-time offenders were apparently lumped with lifetime criminals. Intra-district comparisons did not take account of districts that are subdivided, with courthouses in different areas, urban and rural, and different caseloads. The web site did not report tests of statistical significance to determine whether differences among judges in average sentences were greater than would be expected by chance. For example, the largest district ranked in the bottom ten for the least judge-to-judge differences, the Southern District of California, has a high volume of similar cross-border drug importation cases. See Federal Public Defender, Fact Sheet: TRAC Analysis of Variations in Sentencing Misses the Mark, available at http://www.fd.org/docs/latest-news/fact-sheet-on-trac-3-7-12.pdf. See John Gleeson, The Sentencing Commission and Prosecutorial Discretion: The Role of the Courts in Policing Sentence Bargains, 36 Hofstra L. Rev. 639, 656 n.66 (2008) ("Some drug couriers get a four-level downward role adjustment based on the happenstance of being arrested in New York rather than in Miami, just as some illegal immigrants get a three-level fast-track adjustment based on the happenstance of being arrested in Arizona rather than Utah.").

The form from which the Sentencing Commission gleans much of its sentencing data, the SOR, is simplistic. It was initially designed by the Criminal Law Committee of the Judicial Conference, pre-Booker, to reflect the information that the Sentencing Commission needed to monitor Guideline enforcement, principally departures. (These are the statistics that enable what I have called the Commission's role as "sentencing police, enforcing Guidelines however flawed, rather than acting as real sentencing expert.") n9 Under an agreement between the Sentencing Commission and the Administrative Office of the U.S. Courts (AO), the Commission received information from the AO that would identify the judge in any specific case, but was barred from publicly releasing that data. n10 The PROTECT Act n11 (made unconstitutional by Booker), led to changes in the form as well as the policy on the release of judge identifiers. In an earlier version of the Act, judge identifier data was to be made public--not to enable a sophisticated discussion of sentencing, as in my perfect world, but rather to "out" judges who were departing too much; it was punishment, not public policy. The final version of the Act made judge identifiers available to the Department of Justice and Congress under certain circumstances. n12 The SOR form was then amended to capture government-sponsored departures and post-Booker variances, ostensibly addressing the criticisms of federal judges that had led to the PROTECT Act. n13 After Booker, there were more changes--mainly concerned with capturing how often judges were "varying" (the post-Booker term for a sentence outside the Guideline framework). "Why" remained unaddressed: What about the

Guideline categories were inadequate? What factors were the judges considering that they could not have considered before?

n9 Gertner, supra note 4, at 263.
n10 The agreement states that "[n]o information that will identify an individual defendant or other person identified in the sentencing information will be disclosed to persons or entities outside of the Commission without the express permission of the court for which the information was prepared." U.S. Sentencing Commission, Public Access to Sentencing Commission Documents and Data, 54 Fed. Reg. 51279, 51282 (1989).
n11 Prosecutorial Remedies and Tools Against the Exploitation of Children Today Act (PROTECT Act), Pub. L. No. 108-21, 117 Stat. 650, 667 (2003) (codified as amended at 18 U.S.C. § 3553, 28 U.S.C. §§ 991, 994 (2006)) [hereinafter, PROTECT Act].
n12 See PROTECT Act § 401(h), 117 Stat. at 672. See also Marc L. Miller, Domination & Dissatisfaction: Prosecutors as Sentencers, 56 Stan. L. Rev. 1211, 1245-47 (2004) (describing the information dissemination requirements of the PROTECT Act).
n13 The Department of Justice charged that departure rates went beyond what was contemplated by the Sentencing Reform Act, exceeding 18 percent. A later report demonstrated that increases in departure rates derived not from judges but from prosecutor-supported departures. See U.S. Sentencing Commission, Report to Congress: Downward Departures from the Federal Sentencing Guidelines, at B-28 n.173 (Oct. 2003), available at
http://www.ussc.gov/Legislative_and_Public_Affairs/Congressional_Testimony_and_Reports/Departures/200310_RtC_Downwa
As a result of that finding, the statement of reasons form was amended to be able to account for prosecutor-supported departures, cooperation departures, and judge initiated departures.

For the most part, the SOR form tracks the Guideline categories--the offense level, criminal history--summarized in a section entitled "Court Determination of Advisory Guideline Range (Before Departures)." It requires the judge to check off whether she was sentencing within the advisory range (the Guideline sentence), departing from that range according to "reasons authorized by the sentencing guidelines manual" (an authorized Guideline departure), or sentencing outside the Guideline range (a Booker variance). The Guideline departure categories are listed with the corresponding Guideline sections to be marked.

If the judge decides to sentence outside the advisory range, an important moment in a post-Booker regime, the form is considerably less specific, and surely less useful. First, there are three categories, which continue to reflect the post-PROTECT Act focus on identifying variances with which the government agreed or was at least complicit. Was the sentence the result of (1) plea agreement: binding plea, plea agreement for a variance to which both sides agree, plea agreement in which the government indicates it would not oppose the defense variance motion; (2) a motion for a variance: by the government, by the defense to which the government did not object, or by the defense to which the government did object; or (3) other, presumably a sentence based on a factor the judge alone found to have been significant?

But the next SOR section that could finally identify the reasons for a non-Guideline sentence, that could offer a meaningful window into post-Booker sentencing, however, is even more perfunctory: There are seven very, very general categories largely copied from the 18 U.S.C. 3553(a), like "nature and circumstances of the offense and the history and characteristics of the defendant pursuant to 18 U.S.C. 3553(a)(1)." And while there is an "other" category, a space for the judge to "explain the facts justifying the sentence outside the range," the form as a whole invites a totally formulaic response. Steven Chanenson was spot on: "Ironically, described as a 'Statement of Reasons,' this document contains a parade of nearly meaningless check boxes that mirror the broad § 3553(a) factors. While it does provide space for a factual justification, the form almost seems designed to encourage the kind of mechanical--and arguably unreasoned--approach to sentencing Booker tried to extinguish." n14

n14 Steven L. Chanenson, Write On!, 115 Yale L.J. Pocket Part 146, 147 (2006), available at http://yalelawjournal.org/the-yale-law-journal-pocket-part/criminal-law-and-sentencing/write-on!.

Even the judges who are trying to comply with Booker and articulate non-Guideline reasons for their sentence rarely bother to fill out the form in any detailed way. In too many sessions the SOR is filled out by the courtroom deputy, or the probation officer, figuring out what the judge did or said after the fact. And for some judges, that is a particular challenge. One persists in saying only that his sentence reflected a variance "for all the reasons in 18 U.S.C. 3553(a)," nothing more, making it virtually impossible to identify what factors were actually operative, why the Guidelines were inadequate to accomplish a fair sentence, and much less how this decision might serve as a framework for other judges.

For the judges who do more, who attach a transcript of the sentencing or a sentencing opinion to the SOR, the gesture may be futile for two reasons. First, the Commission's staff codes the data, in effect, recharacterizing it, raising the likelihood that the judge's post-Booker factors will simply be collapsed into the available Guideline categories and that legitimate legal and factual issues will be blurred. When I interpreted "extraordinary family circumstances" or "aberrant behavior" in a fashion that accommodated the purposes of sentencing, rather than according to the strict formalism of the pre-Booker law, was that a Guideline departure or a variance? n15 When I focused on intra-defendant disparity, a non-Guideline category, how was that "coded"? n16 If I departed on a Gall-like ground--age, whether the defendant left the drug trade long before the investigation surfaced--how did the Commission staff deal with it? Or if I sentenced based on the defendant's mental history, without intoning the specific Guideline categories because they did not precisely apply, or put more emphasis on the defendant's minor role than did the Guidelines, how was that categorized? n17 Small wonder that the Commission was surprised that there were sentencing variations for which their data cannot account, or that there were unexplained variations that they suggested may be attributable to creeping racial bias, a vigorously disputed contention. n18 Nothing about the Commission's data collection practices suggests that they cared about the real reasons for the sentencing variances.

n15 See United States v. Germosen, 473 F. Supp. 2d 221 (2007).
n16 See United States v. Haynes, 557 F. Supp. 2d 200 (D. Mass. 2008) (table of codefendants supplied); United States v. Garrison, 560 F. Supp. 2d 83 (D. Mass. 2008) (comparing the sentences of men picked up in the same drug sweep, the same geographical area, with the same charges--including sentences imposed by other judges.); United States v. Whigham, 754 F. Supp. 2d 239 (2010) (same).
n17 See United States v. Whigham, supra note 16 (sentence based in part on a number of factors, including Whigham's severe mental deficits, function at the first grade level though 46 years old.)
n18 Compare, e.g., U.S. Sentencing Commission, Demographic Differences in Federal Sentencing Practices: An Update of the Booker Report's Multivariate Regression Analysis, 2010, available at http://www.ussc.gov/Research/Research_Publications/2010/20100311_Multivariate_Regression_Analysis_Report.pdf, with Jeffrey T. Ulmer et al., Racial Disparity in the Wake of the Booker/Fanfan Decision: An Alternative Analysis to the USSC's 2010 Report, 10 Criminology & Pub. Pol'y 1077, 1086 (2011).

Second, with respect to researchers outside the Commission, like the TRAC researchers, a more detailed SOR would not make much of a difference. All other jurisdictions, apart from Massachusetts, follow the Judicial Conference's troubling recommendation n19 that the SOR be confidential. Small wonder then that TRAC found "hard-to-explain" differences between the median sentences of the judges they studied. Whereas the Commission has access to the SORs but does not bother to capture the reasons for the Booker variances (assuming the forms are carefully filled out), TRAC does not have the reasons at all.

n19 There is at least an argument that keeping the SOR confidential raises First Amendment issues.

Even in Massachusetts, where the SORs are public, the perfunctory nature of the form makes study difficult. Although Professor Ryan Scott, who studied Massachusetts sentencing to evaluate inter-judge sentencing disparity, found troubling signs of its reemergence, his conclusions were necessarily flawed by the flawed data. n20 Unwarranted disparity, after all, is the "different treatment of individual offenders who are similar in relevant ways, or similar treatment of individual offenders who differ in characteristics that are relevant to the purposes of sentencing." n21 But if the forms (and how judges fill them out) do not fully capture the "characteristics that are relevant to the purposes of sentencing," how can the aggregate data do so? So Professor Scott, in effect, monitored judicial "compliance" with the Guidelines as one measure of disparity--important, to be sure, but hardly the complete picture.

n20 Ryan W. Scott, Inter-Judge Sentencing Disparity After Booker: A First Look, 63 Stan. L. Rev. 1 (2010).
n21 U.S. Sentencing Commission, Fifteen Years of Guidelines Sentencing: An Assessment of How Well the Federal Criminal Justice System Is Achieving the Goals of Sentencing Reform, Ch.4, 113 (2004), available at http://www.ussc.gov/Research/Research_Projects/Miscellaneous/15_Year_Study/index.cfm.

Of course, sentences will be disparate if judges approach Booker differently--if there are Guideline judges and Booker judges--but these differences are not capricious or illegitimate. They may well reflect real, ongoing jurisprudential disputes, not personal predilections. Post-Booker judges, after all, are supposed to sentence outside the Guideline range when the range is greater than necessary or insufficient to meet the purposes of sentencing. Indeed, I have argued that there are circumstances when an outside-the-Guideline sentence may not be optional; it may well be essential to prevent both unwarranted disparity and unwarranted uniformity. One way to narrow this kind of disparity would be for the appellate courts to reverse within-Guideline sentences as often as they do outside-the-Guidelines sentences.

So this is not a perfect world. Not merely flawed statistics on the one hand, but hardly a rational discussion on the other. Sentencing debates take place in the supercharged atmosphere of "tough on crime" politicking. The Guidelines were in part the product of a particularly virulent anti-judge atmosphere. During the period when the Federal Sentencing Guidelines were debated, more than two dozen bills stripping or altering federal courts' jurisdiction were introduced. The anti-judge and, significantly, anti-federal judge language of the debate was vituperative. n22 And that atmosphere persists. In the New York Times article reporting the TRAC data, the author was mainly concerned with identifying the judges who were the most lenient or the most strict. Likewise the Scott report was "spun" in the media for identifying which judges were "tough" and which were not. n23 No nuanced account of Massachusetts sentencing; no fair discussion of legal disputes about Booker.

n22 Judge Nancy Gertner, A Short History of American Sentencing: Too Little Law, Too Much Law, or Just Right, 100 J. Crim. L. & Criminology 691, 699 nn.46, 47 (2010). Senator Jesse Helms of North Carolina announced: "Mr. President, unrestricted power has always been the mortal enemy of the rule of law. In our day, we have learned that this is as true in the case of judges as it has been in the case of tyrannical kings and communist dictators. . . . Federal judges have abdicated their role as upholders of the rule of law and become instead tyrants who substitute their own personal views for law. . . ." Christopher LeConey, Rhetorical Branding of Judges as Outlaws: Recasting the SRA of 1984 as Symptom of the Reagan-Era Anti-Judiciary Zeitgeist (on file with author).
n23 Jonathan Saltzman, Disparity Cited in Sentence Lengths, Boston Globe, Dec. 20, 2010 (quoting Ryan Scott as saying that the "findings are troubling . . . because they raise the specter of defendants getting markedly different punishments depending on the politics and biases of the judges before whom they appear." Nothing in the article suggested that the differences were in part the product of jurisprudential differences in the wake of

Booker and in part limitations in the data.

I would love to see full and fair statistics published with judge identifiers--in that perfect world. When I was on the bench, I tried to create a sentencing information system on my own: I asked Probation to give me information about colleagues' average sentences, as well as the bases of their decisions. I would look at sentences for similar offenders. If the SOR gave no clue as to the bases for the judge's decision, I would read opinions, transcripts, even presentence reports, sharing them with counsel with the permission of the parties, or sanitizing them so that they can be distributed to all concerned. I would try to understand why Mr. X in another courtroom got one sentence for a given crime and Ms. Y in my courtroom got another.

In my new academic role, I am continuing that process: I am digitizing all of my sentencing files so that I or someone else can analyze them: my consistency, biases, and so forth, and more significantly, what worked and what did not. And I will publish the results.

But then again, in September of 2011, I retired from the federal bench to join the Harvard Law School faculty. Criticism did not bother me when I was a judge. It surely does not bother me now. A perfect world, this is not.

********** Print Completed **********

Time of Request: Wednesday, February 27, 2013 13:51:29 EST

Print Number: 2828:396734949
Number of Lines: 268
Number of Pages: 7

Send To: Gertner, Nancy
HARVARD UNIVERSITY LAW SCHOOL
1511 MASSACHUSETTS AVE
CAMBRIDGE, MA 02138-2996