Challenging Sentencing Guidelines With Comparable Cases

By **Michael Yaeger** (January 29, 2021)

Michael Yaeger

In December, the U.S. Sentencing Commission issued a report on the influence of the U.S. sentencing guidelines on average sentences imposed. The commission found that the guidelines "generally continue to have a substantial influence on sentences."[1] And influence is the right word.

The ranges of prison time recommended by the sentencing guidelines aren't mandatory any more — but they don't have to be to affect a judge's choice of sentence. The mere act of calculating and considering a guidelines range is enough to anchor a judge to it. And since federal judges are still required to do that calculation, the anchoring effect ripples across the nation.

The question for defense lawyers is how to lift the anchor — how to persuade judges to sentence below the range of months recommended by the guidelines.

Most practitioners answer by spotlighting sympathetic aspects of the defendant's life, which makes sense. But they could enhance their advocacy further with two other resources: statistical analysis of actual sentences, and specific precedents hidden from standard legal search engines.

The commission's latest report, like all of its statistical reports, is not terribly useful for judges deciding what sentence to impose in a particular case, because it is far too general. The commission's reports usually discuss offenses in categories — such as drug offenses or firearms offenses — as opposed to specific Chapter 2 guidelines, such as Section 2D1.1.

Moreover, they almost never provide statistics on particular positions on the sentencing table — i.e., particular recommended ranges of months — which is the level at which the guidelines are applied in individual cases. But the commission's datasets allow for analysis at the level of a particular position on the table, and that can be very valuable.

Many lawyers who commission that kind of statistical analysis — who have the sentencing dataset of more than 1.2 million sentences analyzed for them — will find that people similarly situated to their client actually received sentences significantly below the guidelines range. And a lawyer who can show that fact can give a sentencing judge good reason to let go of the guidelines range.

A discussion of individual, identified, comparable cases can make the point stronger still. Judges presented with the details of comparable cases where a nonguidelines sentence was imposed will have yet another reason to free their thinking.

Such cases are effectively hidden from standard legal search engines, because judges rarely write opinions when they issue sentences. But by reidentifying hundreds of thousands of sentences anonymized in the commission's dataset, we can find the underlying cases. Once found, they can be mined for effective arguments by analogy.

The guidelines have an anchoring effect beyond their legal weight.

An anchoring effect is the influence that a suggested number can have on a person

answering a question or making a decision.[2] Again and again, psychologists and economists have confirmed the existence of anchoring effects experimentally, even with numbers that are obviously chosen at random.

For example, in one experiment, people were asked to spin a wheel with slots numbered zero to 100, then to estimate the percentage of African countries in the United Nations; the ones who spun a higher number on the wheel tended to guess a higher percentage.[3] In another example, real estate agents who were asked to assess the value of a house on the market gave higher valuations when they were told a higher asking price.[4]

Most relevant for our purposes, similar anchoring effects have been observed on experienced judges. One especially striking study tested "German judges with an average of more than 15 years on the bench."[5]

First, the judges read "a description of a woman who had been caught shoplifting."[6] Second, the judges "rolled a pair of dice that were loaded so every roll resulted in either a 3 or a 9."[7] Third, "the judges were asked whether they would sentence the woman to a term in prison greater or lesser, in months, than the number showing on the dice."[8]

Higher rolls of the dice lead to higher sentences. On average, judges who rolled a 9 chose a sentence of eight months, while judges who rolled a 3 chose a sentence of just five months. As explained by Daniel Kahneman, a psychologist and Nobel laureate in economics, that meant the anchoring effect on the judges was 50%.[9]

In the U.S. federal system we don't use loaded dice, but we're required to calculate the sentencing guidelines — and they appear to have a similar effect. According to the commission's 2019 annual report and sourcebook of federal sentencing statistics, there were 76,034 sentences imposed in fiscal year 2019, and over 51% of them were within the guidelines range (another 21% were departures on a government motion).[10]

Judges don't merely consider the guidelines range: They tend to impose it. And the experimental literature touched on above suggests that the influence of a guidelines range may extend even to a judge who is not consciously adhering to it.

The guidelines are a subtle opponent that benefits from the status quo. To combat them, defense counsel must provide judges with ways to slow down, step back from the well-worn path of the guidelines, and engage their conscious, more deliberative faculties so that they can make an active and free choice of a sentencing range.

The guidelines' sentencing ranges are vulnerable to empirical attack.

Statistics on the sentences that judges actually impose can counteract the effect of the guidelines recommendations — not just because statistics offer rival numbers, but also because the theoretical underpinning of the guidelines make them especially vulnerable to empirical attack.

At their inception, the guidelines based their claim to authority on their purported empiricism. The Sentencing Reform Act of 1984 directed that the commission "establish sentencing policies and practices that … assure the meeting of the purposes of sentencing as set forth in section 3553(a)(2)."[11]

However, Congress gave no direction on which purposes should predominate — for example, on whether principles of just deserts should be given greater weight than

incapacitation — so the commission sidestepped the issue. It claimed that the guidelines it issued derived principles and sentence lengths from empirical study of past sentences.[12]

Moreover, the U.S. Supreme Court has made clear that the authority the guidelines have — or lack — depends on how empirical they are. Specifically, the court held in its 2007 decision in Kimbrough v. U.S. that when the commission diverges from "empirical data and national experience," it diverges from its "characteristic institutional role," and thus has less persuasive authority than it otherwise would.

Accordingly, the guidelines' claim to authority can be weakened by showing that they do not accord with actual sentencing practice. Detailed yet straightforward statistics that answer the guidelines' recommendations on their own terms can provide the basis for a judge to consciously resist the anchoring effect of the guidelines' ranges.

Counsel's empirical arguments should be as granular as the sentencing table.

None of the commission's reports — including its most recent report on the guidelines' influence — provides comparisons of defendants within the same Chapter 2 offense guideline, offense level and criminal history category. But the commission's data files make such an analysis possible.

We can compare, for example, defendants sentenced under guidelines Section 2B1.1, offense level 20 and criminal history category I. Those defendants reside in precisely the same position on the sentencing table, so they face the same recommended range of imprisonment, namely, 33-41 months.

Defense lawyers should commission analyses of those cohorts to see how many of these similarly situated defendants receive guidelines sentences. If, as is true for many cohorts, the median sentence is under the recommended guidelines range, that is a powerful response to the government's push for a guidelines sentence.

After all, a guidelines sentence is one within the specific range of months where the applicable guideline, offense level and criminal history category place a defendant. When the government seeks a "guidelines sentence" for a defendant, it is implicitly asserting that he or she deserves the same sentence given to other people who share the same position on the sentencing table.

If 50% of those people are receiving sentences beneath the guidelines range, that seriously undermines the claim that a guidelines sentence will prevent unwarranted disparities between similarly situated defendants. To the contrary, in that situation a guidelines sentence may create such disparities.

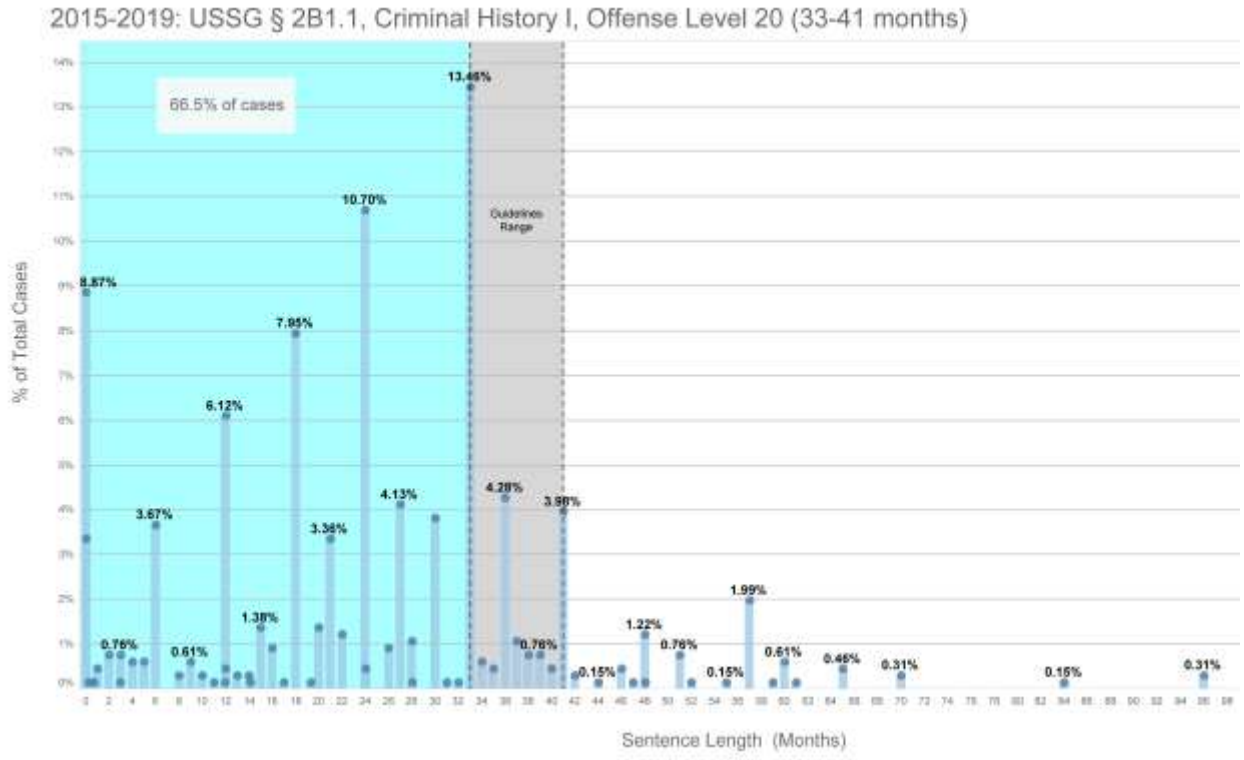
A quick case study shows how empirical analysis can help lift the guidelines' anchor.

The guidelines range for the example above — Section 2B1.1, offense level 20, criminal history category I — is 33-41 months, but the actual median sentence is far lower.

Our research at Empirical Justice found that, from 2015 to 2019, there were more than 1,000 defendants in that cohort, and the actual median sentence was 24 months. Not the average, the median.

In other words, even though the guidelines recommended a sentence of 33-41 months,

50% of those defendants received a sentence of 24 months or less. That's nine months less (and 25% lower) than the bottom of the guidelines range. Moreover, over 66% of the defendants received sentences under the guidelines range. The histogram below illustrates these figures visually and also provides additional detail:



Also, it's worth observing that the most common sentence is 33 months — the precise bottom of the guidelines range — and the third most common sentence is probation. Perhaps the judges who imposed 33 months thought they were imposing a relatively low sentence. And perhaps if those judges knew that they were actually imposing a punishment higher than the sentence in 66% of similar cases, they would have sentenced below the guidelines range, not inside it.

A prosecutor seeking a guidelines sentence — and, not incidentally, trying to bolster the view that judges should sentence under the guidelines only in response to a government motion — might object to the analysis above.

For example, the prosecutor might object that the statistics above include cooperating witnesses receiving credit for "substantially assisting in the prosecution of another" under guidelines Section 5K1.1. Putting aside the question of exactly how differently cooperators should be treated from other defendants, prosecution and defense alike may be surprised by the results when cooperators are removed.

In fact, when cooperators are excised from the sample, 50% of the defendants are still sentenced to just 27 months, and over 58% still receive sentences beneath the bottom of the guidelines range — i.e., below 33 months.

In sum, statistical analysis of actual sentences can help defense counsel show that a

sentence substantially below the applicable guidelines range is not exceptional, but typical. It gives judges reason to doubt that the range is rooted in "empirical data and national experience."

Hidden sentencing precedent can also deter reflexive adherence to the guidelines.

Statistical analysis of actual sentences can be even more valuable when coupled with individual, identified cases of similarly situated defendants who have received nonguidelines sentences. But the cases are hard to find. Standard legal search engines aren't much help because most sentences are imposed without a written opinion.

And the commission deliberately deidentifies cases in its data files. It strips out the names of judges and cases, it removes cases numbers, it creates completely new identification numbers, and it cuts out the individual day on which a defendant is sentenced, leaving just the month and year.

But there are ways to reidentify cases in the Sentencing Commission's dataset. And doing so allows lawyers to couple their statistical attack with a detailed discussion of particular apposite cases.

When both stats and identified cases show that defendants in a particular range are receiving nonguidelines sentences, the guidelines are unmoored from their authority. They are revealed as something other than empirical.

Michael L. Yaeger is a shareholder at Carlton Fields and a former federal prosecutor for the Eastern District of New York. He is the founder of Empirical Justice LLC.

[1] U.S. Sentencing Commission, The Influence of the Guidelines on Federal Sentencing: Federal Sentencing Outcomes, 2005-2017, https://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publications/2020/20201214_Guidelines-Influence-Report.pdf#page=9.

[2] See Daniel Kahneman, Thinking, Fast and Slow 120-128 (2011); see also Adrian Furnham and Hua Chu Boo, A literature review of the anchoring effect, J. Socio-Economics, Volume 40, Issue 1, Feb. 2011, 35-42, available at https://www.communicationcache.com/uploads/1/0/8/8/10887248/a_literature_review_of_the_anchoring_effect.pdf.

[3] Kahneman at 120.

[4] Id. at 124.

[5] Id. at 125.

[6] Id.

[7] Id.

[8] Id. at 125-26.

[9] Id. at 126. The difference between the dice rolls is 6, and the difference between the average sentences was 3 months, so the ratio of the two differences (3/6) expressed as a percentage is 50%.

[10] U.S. Sentencing Commission, 2019 Annual Report and Sourcebook of Federal Sentencing Statistics, https://www.ussc.gov/sites/default/files/pdf/research-and-publications/annual-reports-and-sourcebooks/2019/Table29.pdf.

[11] 28 U.S.C. § 991(b)(1).

[12] See U.S.S.G., Ch. 1 pt. A(3) (1987); Justice Stephen Breyer, The Federal Sentencing Guidelines and the Key Compromises Upon Which They Rest, 17 Hofstra L. Rev. 1, 7 (1988); Justice Stephen Breyer, Justice Breyer: Federal Sentencing Guidelines Revisited, 11 Fed. Sent. R. 180, 1999 WL 730985 *3 (Vera Inst. Just.) (Feb. 1, 1999).