Sentencing in Chaos: How Statistics Can Harmonize the "Discordant Symphony"

MARK H. ALLENBAUGH*

Chief Research Officer, Sentencing Stats LLC

I. Introduction

Prior to the creation of the U.S. Sentencing Guidelines, federal judges had nearly unfettered discretion to impose sentences anywhere they deemed appropriate within the applicable statutory ranges.[1] This lack of restraint, not surprisingly, led to "law without order," as Judge Marvin Frankel famously observed.[2] Sentences imposed on similarly situated defendants varied dramatically from judge to judge, as each judge weighed different factors differently under a variety of competing penological philosophies. That parole, too, was possible added significant uncertainty to an already widely disparate sentencing landscape.

In response, Congress passed the Sentencing Reform Act of 1984, which created the U.S. Sentencing Commission ("Commission").[3] The Commission, in turn, was tasked with drafting the U.S. Sentencing Guidelines ("guidelines"), which were intended to bring rationality to federal sentencing by promoting uniformity, proportionality, and certainty.[4] Uniformity was to be achieved by ensuring that similarly situated offenders were sentenced similarly across the country. Proportionality was to be achieved by carefully identifying and weighing particular factors relevant to the seriousness of the offense conduct and the offender's culpability, and calibrating those to a specific, narrow sentencing range that was no more than six months to a few years in length. Certainty was to be achieved primarily by eliminating parole and ensuring that at least 85% of any sentence imposed was actually served, but it was also to be achieved by setting mandatory and narrow sentencing ranges that corresponded to the Total Offense Level (TOL) and Criminal History Category (CHC) of a defendant.

When first promulgated on November 1, 1986, the guidelines were mandatory.[5] As such, and as originally intended by Congress, there was very little discretion for judges to exercise at sentencing. Once the guidelines were calculated and a sentencing range thereby determined, a judge was required to sentence within that rather narrow range, save for those rare instances when there was an adequate ground to "depart" below or above that range. The individual guidelines thus were both descriptive and prescriptive in nature: they described the relevant factors to consider at sentencing and prescribed the particular weight to assign them. Once such weighing was complete, the Sentencing Table[6] then prescribed the approximate sentence to impose by providing a narrow sentencing range. Given that the guidelines were mandatory, the prescribed sentencing range also *described* the approximate sentence imposed for similarly situated offenders.

Thus, while under this mandatory regime judges' sentencing discretion was greatly reduced, uniformity, proportionality, and certainty were rather easy to achieve. Defendants sharing the same guidelines calculation would, for the most part, be prescribed identical sentencing ranges. Save for the "rare" court-initiated departure or governmental motion for a downward departure,[7] such defendants would receive a sentence within that narrow range. At the very least, if there were any systemic issues with sentencing, these were a function of the guidelines, the Commission, and Congress, but not the exercise of judicial sentencing discretion. The Commission simply would amend the guidelines when it determined that changes were needed or when Congress so directed.[8]

All this fundamentally changed on January 12, 2005. In a still unique double-majority opinion, the U.S. Supreme Court in *United States v. Booker*[9] held that (1) mandatory guidelines violated the Sixth Amendment[10] and, as a result, (2) the guidelines were now to be treated as "merely advisory."[11] As such, while judges still had to calculate the guidelines correctly to determine the TOL and CHC, they need only "consider" the guidelines' sentencing range. Judges could once again impose any sentence they wanted (within statutory constraints, of course).

Dissenting from the remedial opinion, Justice Scalia presciently observed that a merely advisory sentencing scheme would bring a "discordant symphony" to federal sentencing:[12]

> What I anticipate will happen is that "unreasonableness" review will produce a discordant symphony of different standards, varying from court to court and judge to judge, giving the lie to the remedial majority's sanguine claim that "no feature" of its avant-garde Guidelines system will "ten[d] to hinder" the avoidance of "excessive sentencing disparities."[13]

As reviewed in detail below, he could not have been more correct.

Fifteen years after the *Booker* decision, federal sentencing is far from uniform and is growing more discordant. Sentences are being imposed at increasing

Federal Sentencing Reporter, Vol. 32, No. 3, pp. 128–137, ISSN 1053-9867, electronic ISSN 1533-8363.
 DOI: https://doi.org/10.1525/fsr.2020.32.3.128.

rates and degrees of variance from the guidelines, both individually and on the whole. While federal offenders still are generally required to serve at least 85% of their sentences, the wide disparity in sentences imposed has, necessarily, injected a high level of uncertainty in federal sentencing.

But this is not to say that all is lost nor that the guidelines no longer serve any purpose. The guidelines can and should be saved. This article argues that by considering data and statistics of sentencing *distributions* imposed on both similarly and dissimilarly situated offenders, federal sentencing can be harmonized. The guidelines, after all, were born of data and statistics, and this article argues that they can be saved by them as well.

Part II discusses the impact of *United States v. Booker* on federal sentencing and provides some graphic comparisons of pre-*Booker* and recent sentencing trends. The difference in sentencing trends is dramatic.

Part III corrects a common misperception: that the guidelines as a whole are merely advisory. As discussed, the guidelines in fact remain mandatory, and any incorrect calculation constitutes reversible error unless the error was harmless (a rarity). Rather, it is the Sentencing Table that is merely advisory. Correctly calculating the guidelines is still required, so the TOL and CHC (as well as other pertinent factors) provide a straightforward way to identify similarly situated individuals, something that is otherwise undefined by statute or the guidelines. Analyzing the sentencing distribution of similarly situated individuals so defined can then provide a suitable range for imposing a sentence. In other words, the mandatory guidelines' calculation provides a method for identifying similarly situated offenders, and the Commission's sentencing data identify an appropriate range for imposing a sentence. The advisory Sentencing Table is, in effect, functionally useless, as several statistical charts illustrate.

Part IV discusses several upshots of this empirical approach, which effectively creates a digital common law of federal sentencing. This approach ultimately is the most effective method for harmonizing the primary purposes of the guidelines with real-world sentencing practice.

Finally, Part V concludes by arguing that, rather than resort to the guidelines' Sentencing Table as the starting point to consider when imposing a sentence, courts should instead reference the Interquartile Range—a statistic commonly used to describe the central distribution of data—of sentences imposed on those similarly situated. Moreover, the U.S. Sentencing Commission should recalibrate the Sentencing Table downward so as to match actual sentencing practice.

II. The Symphony Becomes Discordant: *United States v. Booker*

The guidelines were developed to achieve three primary penological goals: uniformity, proportionality, and certainty in sentencing.[14] In essence, uniformity was achieved by sentencing similarly situated offenders similarly; proportionality was achieved by ensuring that the severity of the sentence was proportional to the seriousness of the offense conduct; and certainty was achieved in two ways: first by ensuring that the bulk of the sentence imposed was the sentence actually served, and second by providing a relatively narrow sentencing range. The sentencing range effectively provided notice to the defendant of what his likely sentence would be, based on the guidelines calculation.

The goals of uniformity, proportionality, and certainty could be achieved only by designing the guidelines, *including the Sentencing Table*, to be mandatory. To be sure, the Commission recognized that there would be cases presenting unusual circumstances that could warrant a departure from the guidelines. But those would be "rare occurrences."[15]

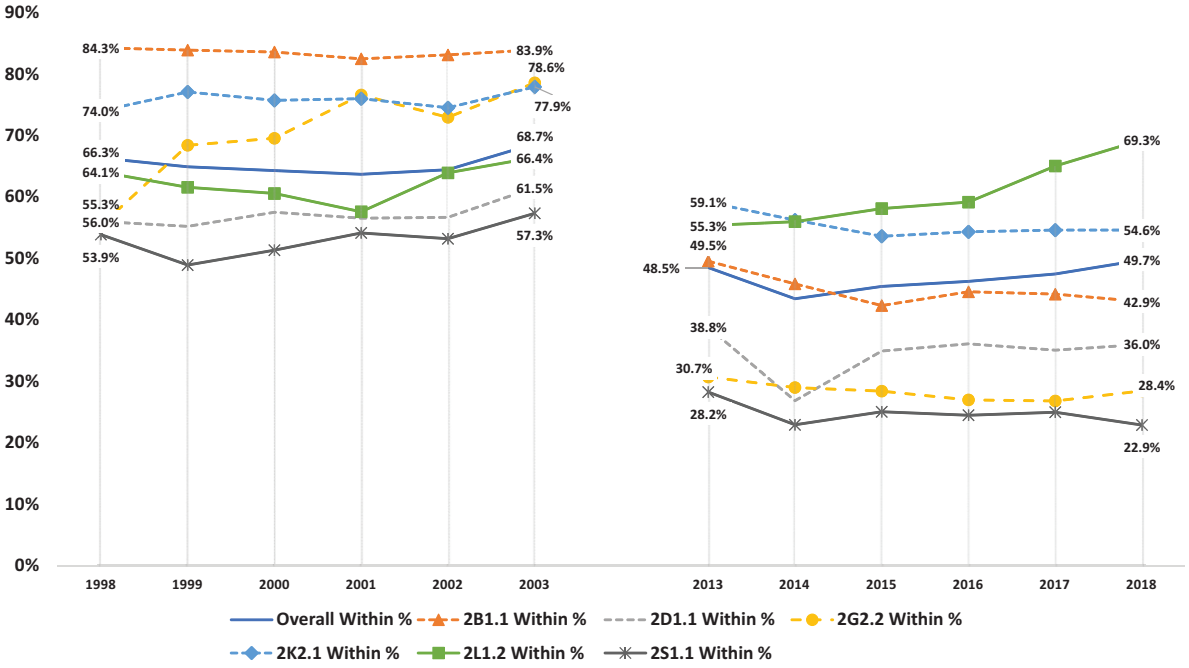
On January 12, 2005, the floor fell out of the guidelines and the ceiling was removed. The U.S. Supreme Court held in *United States v. Booker*[16] that, to comply with the Sixth Amendment, the guidelines had to be treated as "merely advisory."[17] While courts were still required "to consider guidelines ranges," courts were now permitted "to tailor the sentence in light of other statutory concerns as well."[18] *Booker* thus fundamentally changed sentencing practice.

Figure 1 illustrates just how starkly *Booker* changed federal sentencing practice. It compares the frequency of within-guidelines sentences for select guidelines during the six-year period prior to *Booker* and its precursor *Blakely v. Washington*[19] and during the most recent six-year period, as reported by the Commission. The overall compliance rate—meaning sentences within the applicable range as defined by the Sentencing Table—dropped from over 68% prior to *Booker* to less than 50% currently. However, the change for specific guidelines has been even more dramatic. For example, in the six-year period prior to *Booker*, the compliance rate for sentences imposed under USSG § 2G2.2 (the guideline covering child pornography possession and distribution offenses) rose from 55.3% in 1998 to 78.6% in 2003; in stark contrast, over the most recent six-year period, compliance rates dropped from 30.7% in 2013 to 28.4% in 2018. A similar contrast is seen for sentences imposed under USSG § 2B1.1, which covers most fraud offenses: prior to *Booker*, the compliance rate for that guideline remained rather steady at approximately 84%; by 2018, the compliance rate had dropped by nearly half, to 42.9%.

The compliance rates for all the selected guidelines have decreased dramatically since *Booker*, save for USSG § 2L1.2, which covers illegal reentry offenses; in stark contrast to the trends for the other guidelines, compliance rates for such sentences have actually increased post-*Booker*. This is almost certainly because the TOLs for those sentenced under USSG § 2L1.2 tend to be quite low, leaving little or no room for downward departures or variances in any event.

Of course, some of the compliance rates may be affected by prosecutorial discretion in terms of frequency

Figure 1
Rate of Sentences Within Applicable Guidelines Range by Selected Guidelines (1998–2003; 2013–2018)



Source: U.S. Sentencing Comm'n, *Sourcebook of Federal Sentencing Statistics*, tbl. 28

of moving for departures pursuant to USSG § 5K1.1 (departures for substantial assistance) or USSG § 5K3.1 (fast-track departures). Accordingly, Figure 2 illustrates the rate of court-initiated departures and variances under the guidelines.[20] It is clear that courts have, to varying degrees but significantly so, exercised their authority to vary downward from the applicable sentencing range. Overall, courts departed below the guidelines only 16.9% of the time in 2002 but departed or varied below the guidelines nearly 30% of the time as of 2018. Again, nearly all guidelines surveyed saw a significant increase in the rate of departures or variances post-*Booker*, with the exception of USSG § 2L1.2. As reviewed above, the compliance rates with that guideline actually increased post-*Booker*, while the court-initiated departures or variances decreased. The explanation, though, is likely the same: that particular guideline generally produces very low sentencing ranges, such that courts generally cannot vary or depart below that range.

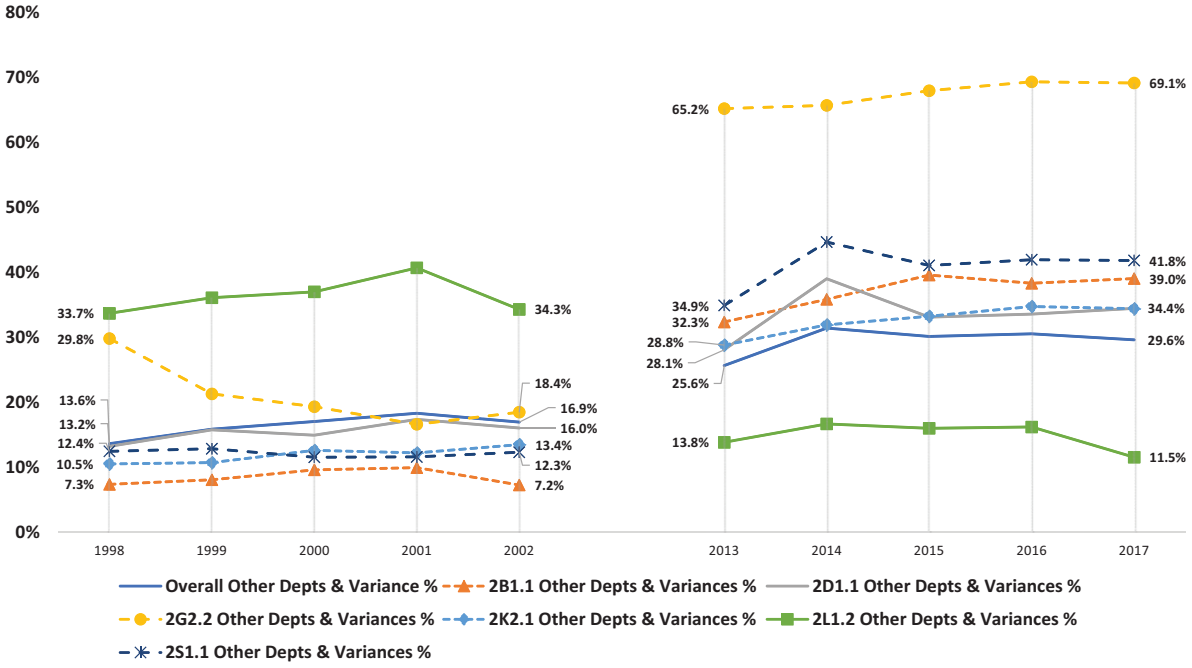
III. Retuning to the Sentencing Symphony: What Does "Similarly Situated" Mean?

It is often said that *Booker* rendered the guidelines advisory, suggesting that the guidelines as a whole are merely advisory. *Booker* did not and they are not. As the Supreme Court has consistently reiterated, the guidelines still must be correctly calculated in every case and an incorrectly calculated TOL or CHC constitutes plain error.[21] And this is so even if the sentencing judge had calculated the incorrect range but coincidentally imposed a sentence within the correct range,[22] or departed or varied from the incorrect range into the correct range.[23] Indeed, cementing the *mandatory* nature of the guidelines calculation is the fact that amendments to the guidelines are still subject to the ex post facto clause.[24]

Booker thus did not result in the guidelines writ large becoming advisory. Rather, what *Booker* made advisory was nothing more than the Sentencing Table (i.e., the list of sentencing ranges set forth in that table resulting from a correct calculation of the guidelines). Understanding this fundamental distinction reveals the path toward saving the guidelines from irrelevancy.

So, if similarly situated individuals are to be sentenced similarly, it is first necessary to determine what constitutes "similarly situated." Curiously, especially given its central function in federal sentencing, "similarly situated" is defined neither by statute nor within the guidelines themselves. Section 3553(a)(6) of Title 18 of the U.S. Code provides that, in imposing sentence, courts are to consider, inter alia, "the need to avoid unwarranted sentence disparities among defendants with similar records who have been found guilty of similar conduct." But what factors does one use to determine whether defendants are similarly situated? The actual offense of conviction? If so, is one convicted of mail fraud similarly situated to one convicted of wire fraud? Tax fraud? Are all economic crimes similarly situated? And what other factors should or should not be considered? What if offenders are of different races or genders or ages? The U.S. Code provides little guidance. As for the guidelines, the phrase "similarly situated" occurs

Figure 2
Rate of Other (Non-5K1.1/5K3.1) Departures or Variances by Selected Guidelines (1998–2002; 2013–2017)



Source: U.S. Sentencing Comm'n, *Sourcebook of Federal Sentencing Statistics*, tbl. 28

only five times throughout the entire Manual: twice at USSG § 1B1.12, discussing the sentencing of juvenile delinquents, and once each in Application Note 4(B) to USSG § 2C1.1, Application Note 3(B) to USSG § 2C1.3, and Application Note 1(B) to USSG § 5H1.7. In none of those instances is the phrase defined.

Intuitively, of course, it would seem most natural to define offenders as similarly situated when they share, at least, the same guidelines calculation (i.e., the same guideline, TOL, and CHC). In such situations, the offenders face the same advisory sentencing range. Indeed, such an approach appears to have been adopted, at least implicitly, by the U.S. Supreme Court.

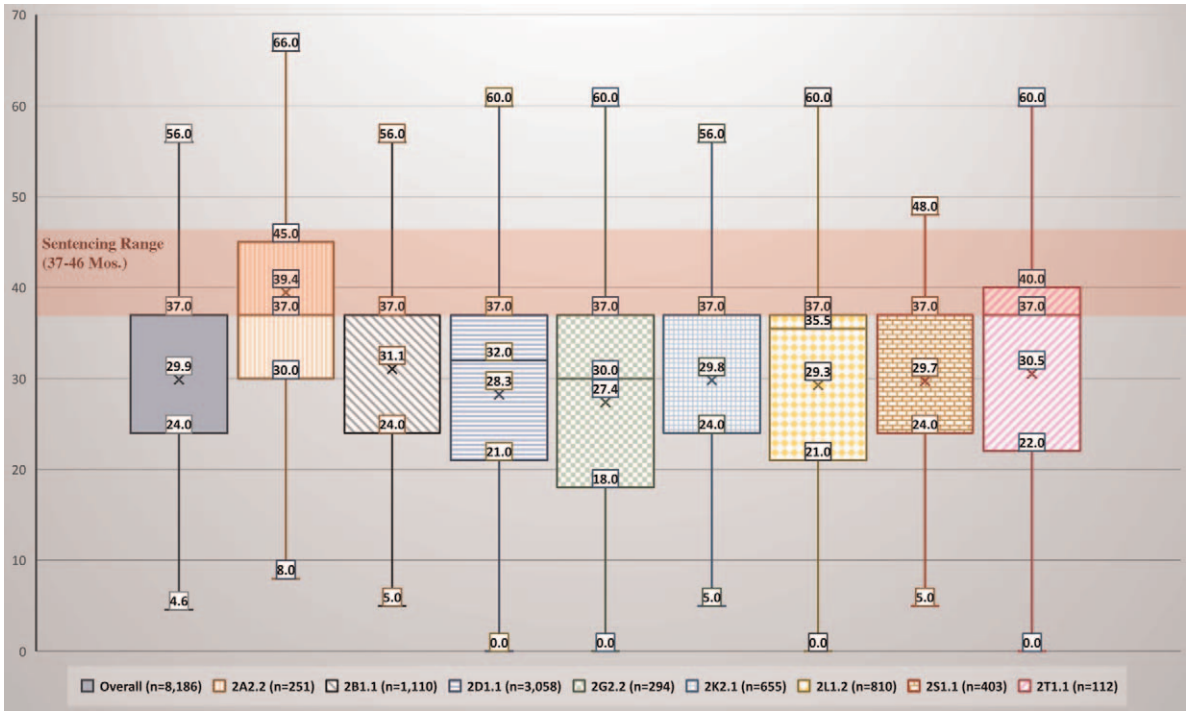
In *Peugh v. United States*,[25] the Court was confronted with whether the ex post facto clause applies to the guidelines, notwithstanding their advisory nature. In holding that it does, the Court, citing *Miller v. Florida*,[26] observed that, like the guidelines, Florida's "sentencing scheme . . . was designed to channel sentences for similarly situated offenders into a specified range."[27] In short, the Court used the "specified range" as the determinative factor for whether offenders were, in fact, similarly situated. Defendants sharing the same "specified range" could thus be considered similarly situated for sentencing purposes.

At least one (now former) sentencing judge, the prolific and scholarly Nancy Gertner, has expressly found the guidelines to take this approach: "The guidelines define 'similarly situated' only with reference to the particular guideline categories. If a defendant had an offense level of 14 and a criminal history of I, the guidelines assumed that you were similarly situated to other 14s and Is."[28] Another (now former) judge, Mark Bennett, also adopted Judge Gertner's definition of "similarly situated."[29]

To be sure, there may be other ways to define "similarly situated," but *for guidelines purposes*, identifying those cases with identical guidelines calculations appears to be the most logical way to proceed. Of course, other pertinent factors should be taken into consideration. For example, those not subject to mandatory minimum or consecutive penalties may not be similarly situated to those who are. In addition to limiting a court's sentencing discretion, a mandatory minimum penalty can alter the guidelines range if it is greater than the otherwise applicable minimum of that range.[30] Likewise, the statutory maximum penalty can alter the guidelines range if it is less than the maximum of the otherwise applicable range.[31] Accordingly, when determining whether cases are similarly situated, in addition to taking into consideration the applicable guideline, the TOL, and the CHC, one must also consider the statutory minimum, statutory maximum, and any mandatory consecutive sentences. Moreover, the Commission itself appears to have adopted this methodology for identifying similarly situated offenders.[32] Thus, in addition to the applicable guideline, TOL, and CHC, any statutory limits on an offender's sentence are key factors to consider when determining whether offenders are similarly situated.[33]

So, with this methodology for identifying what constitutes similarly situated individuals, we can now turn to the

Figure 3
Sentencing Distribution by Selected Guideline TOL 21, CHC I, No Man. Min. or Consec., Stat. Max. 48+ Mos., No 5Ks (2006-2018)



data. In theory, and as reviewed above, offenders with identical guidelines calculations and sentencing ranges should receive sentences within that range. And that was generally the case pre-*Booker*. Now, however, the results are quite different.

Figure 3 compares the sentencing distributions for offenders sentenced under various guidelines between fiscal years 2006 and 2018, where all offenders had been assigned a TOL of 21 in CHC I,[34] and where none were subject to a mandatory minimum or consecutive sentence but where all were subject to a statutory maximum sentence of at least forty-eight months. Thus, all offenders were assigned a sentencing range of thirty-seven to forty-six months, per the guidelines' Sentencing Table.

Box plots (sometimes referred to as box-and-whiskers plots) are a "common statistical tool" utilized by the Commission in its official statistical reports.[35] They provide a visualization of the distribution of sentences and therefore provide more insight into sentencing practice than would a static number alone, such as the average or median.

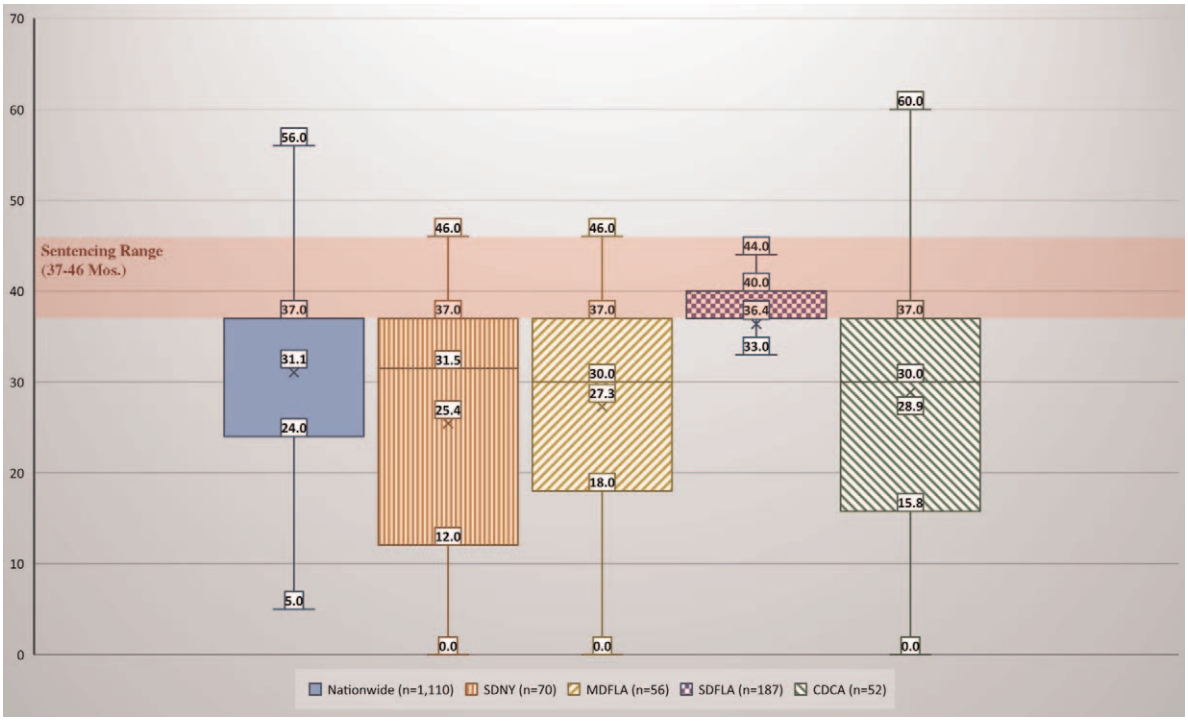
The "x" within the box indicates the average sentence. For example, for the 251 sentenced under USSG § 2A2.2 (aggravated assault), the average sentence was 39.4 months. The line bisecting the box indicates the median, which was thirty-seven months for those sentenced under USSG § 2A2.2. The top of the box indicates the third quartile or 75th percentile, meaning that 75% of those sentenced under USSG § 2A2.2 received a sentence of forty-five months or less. The bottom of the box indicates the first quartile or 25th percentile, which was thirty months in the case of USSG § 2A2.2. Thus, half of all those sentenced under USSG § 2A2.2 and meeting the identified criteria received a sentence between thirty and forty-five months. This box is sometimes referred to as the interquartile spread or distribution.[36]

The "whiskers," the lines extending below and above the boxes, do not necessarily correspond to the maximum or minimum sentences for the identified group (although they can), but rather identify the outlier threshold: any sentences below or above these thresholds are considered statistical outliers or anomalies.[37] For those sentenced under USSG § 2A2.2, for example, the bottom threshold was eight months while the top threshold was sixty-six months. The applicable sentencing range for all offenders is identified by the transparent red rectangle.

Two observations are immediately apparent. First, for nearly all the selected guidelines, the majority of sentences imposed were well below the bottom of the applicable range. Only those sentenced under USSG § 2A2.2 had at least half of their sentences imposed within the applicable range. Second, the distribution of sentences is remarkably different among the various guidelines. While the top of the box is the same for most, the bottom varies considerably, as do the averages and medians for each group of offenders.

The first observation suggests that the sentencing range is calibrated too high. That this is the case for remarkably different offense types—fraud, drug distribution, child pornography, and so on—indicates this is so. Only for

Figure 4
Sentencing Distribution by Selected Districts 2B1.1, TOL 21, CHC I, No Man. Min. or Consec., Stat. Max. 48+ Mos., No 5Ks (2006-2018)



aggravated assault offenses do judges find the range appropriate, and then only half the time.

The second observation suggests that judges are taking the offense type itself into consideration—or, at least, some other factor that is not contained within the guidelines themselves. For example, and rather remarkably, the first quartile for fraud offenses under USSG § 2B1.1 is twenty-four months, whereas it is only eighteen months for child pornography possession and distribution offenses under USSG § 2G2.2. It is beyond the scope of this article to offer a definitive answer to why there are such discrepancies, other than to point out that there are additional, non-guidelines factors that may be specific to the nature of the offense, which courts are taking into consideration when imposing a sentence.

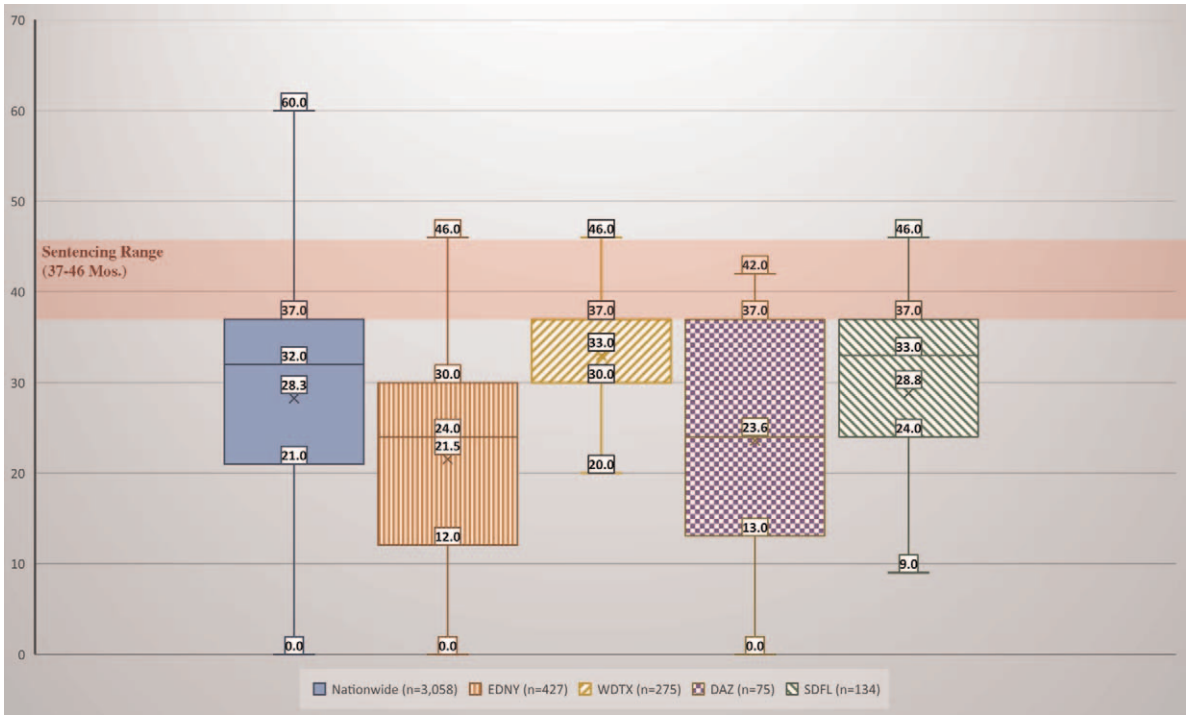
Are there other non-guidelines factors that affect the sentencing distribution? Yes. As the Commission has found, geography can make a substantial difference even though in theory it should not.[38] Figure 4 compares the sentencing distributions for all those sentenced under USSG § 2B1.1 and meeting the same criteria as above. The disparity among the selected districts is rather remarkable—especially the high but narrow distribution for the Southern District of Florida—and strongly suggests not only differences in sentencing philosophy among different judges, but also that charging and plea practices among local U.S. attorneys' offices contribute to sentencing practice. Notably, for three of the four districts surveyed, the majority of sentences imposed were below the guidelines range.

This phenomenon, to be sure, is not specific to fraud offenses sentenced under USSG § 2B1.1 but is also observed for drug trafficking sentences imposed under USSG § 2D1.1, as shown in Figure 5. Again, there was significant disparity among the districts surveyed, and notably, the majority of all sentences were still imposed below the bottom of the applicable sentencing range. There is no reason to believe that similar inter-district disparities would not be found for sentences imposed under other guidelines. Ultimately, the point here is not to definitively identify the causes of such disparities, but rather to illustrate how use of sentencing data can identify where sentences for similarly situated offenders tend to fall.

IV. Harmonizing Sentencing

In light of the advisory nature of the Sentencing Table, federal courts across the country have increasingly sought the guidance of the Commission's sentence statistics when seeking to implement the Congressional directives set forth at 18 U.S.C. § 3553(a). As the U.S. Court of Appeals for the Sixth Circuit recently observed, "national sentencing data released by the Commission should serve as 'a starting point for district judges in their efforts to avoid unwarranted sentence disparities.' "[39] Of course, such data must have "sufficient detail and context to show that the defendants whose sentences were reflected by the statistics 'ha[d] been found guilty of similar conduct.' § 3553(a)(6). . . . [A] sufficient level of fit is required between the cited statistics and the case at hand."[40] "[G]eneral statistics that cover

Figure 5
Sentencing Distribution by Selected District 2D1.1, TOL 21, CHC I, No Man. Min. or Consec., Stat. Max. 48+ Mos., No 5Ks (2006-2018)



a multitude of other crimes committed in a multitude of other ways do not create an 'unwarranted' disparity."[41]

As argued above, utilizing the Commission's data files to identify those offenders with the same guideline, TOL, and CHC, while also taking into account any mandatory minimum or consecutive penalties as well as the statutory maximum provides a sufficient level of "fit" to define the class of "similarly situated" offenders. And while unwarranted disparity is defined with respect to a national perspective,[42] as also noted above, geographic variations should, *in addition*, be taken into consideration inasmuch as they can identify for courts important factors regarding local sentencing and prosecutorial practices that otherwise may be hidden within a nationwide data set.

What is most clear from these analyses and the Commission's own reports is that the Sentencing Table has lost its relevance. Courts, for most offense types, are increasingly imposing sentences below the otherwise applicable sentencing range. And as the offense level increases, the magnitude of the downward variance generally increases too. This is exemplified in Figure 6, which illustrates the rate of court-initiated downward variances by TOL in CHC I.[43] As indicated, there was a 2759% downward variance rate for all offenders with a TOL of nine in CHC I.[44] That rate rose to 58.1% for offenders with a TOL of 30, then fell to the mid-40s before once again rising to 57.7% for offenders with a TOL of 43, which corresponds to life imprisonment,[45] or the overall statutory maximum penalty if life is not available.[46]

Consistent with Figure 6, Figure 7 illustrates the average sentence imposed as the result of a court-initiated downward variance by TOL.[47] As can readily be seen, the gap between the bottom of the applicable sentencing range and the average sentence imposed (that resulted from a downward variance) generally increases as the offense level increases.

V. Fine Tuning the Sentencing Symphony

In light of the increasing rates and degrees of variance, the Sentencing Table is no longer relevant as either a descriptive or a prescriptive matter. Not only do the specific rates and degrees of variance fluctuate dramatically from guideline to guideline at the same TOL and CHC, but they do so even within the same guideline at the same TOL and CHC from district to district. Thus, while courts must continue to properly calculate the guidelines, is such an exercise cynical, given that the goals of uniformity, proportionality, and certainty are not being achieved?

The process of calculating the guidelines, as discussed above, is a sound mechanism for identifying similarly situated offenders from within the Commission's data files. As demonstrated above, once the class of similarly situated offenders is identified by application of the guidelines, pertinent statistics can then readily be gleaned from the Commission's robust data files to determine what actual sentencing practice is. But this does not mean that courts should simply find the average or median

Figure 6
Rate of Downward Variances by Total Offense Level in CHC I Excluding Man. Min., Consec., and 5Ks (2006–2018) (n=329,317)

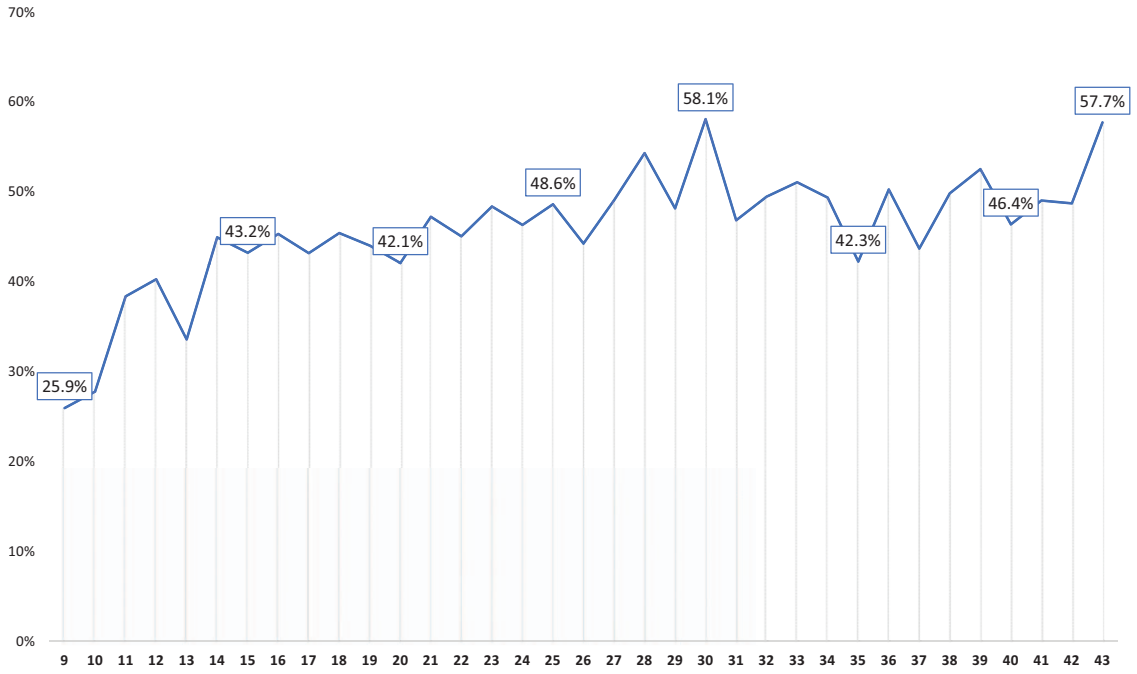
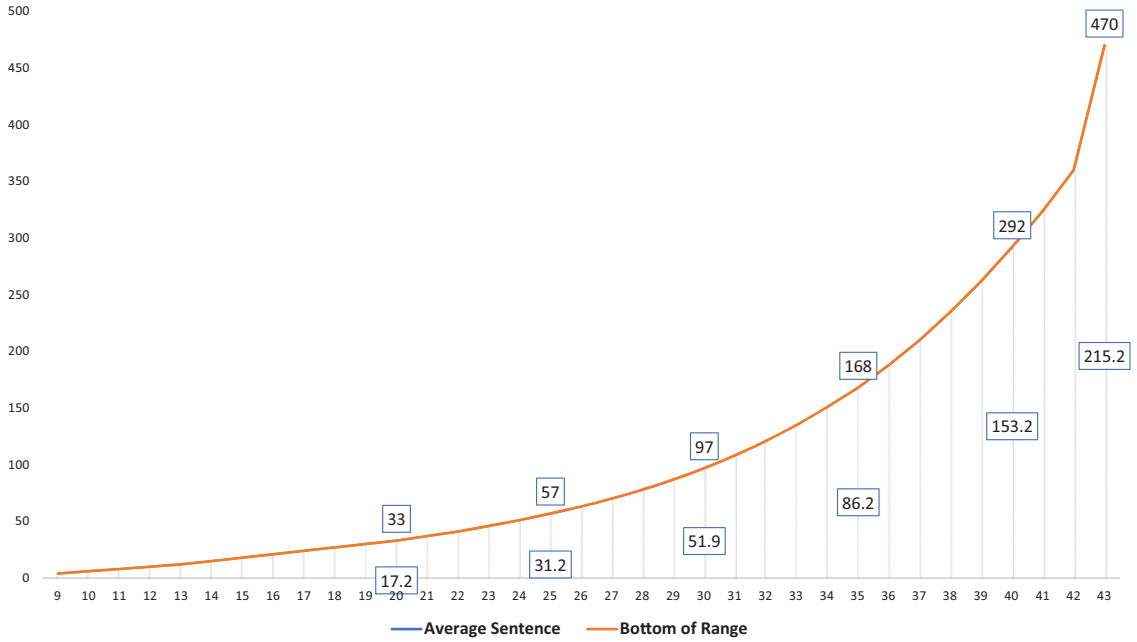


Figure 7
Comparison of Sentences Imposed as a Result of a Downward Variance by Bottom of Offense Level Range Excl. Man. Min., Consec., and 5Ks (2006–2018) (n=61,886)



sentence for those similarly situated and impose that sentence. As the Commission itself recognized when first promulgating the guidelines, there was no practical way to account for every pertinent factor in every case. Hence the necessity of departures.

So, while averages and medians may give a court a good idea of the central tendency of sentences for a particular class of similarly situated offenders, they do not necessarily prescribe what the sentence should be, especially where unusual aggravating or mitigating factors may be present.

Thus, as with the structure of the Sentencing Table, a range rather than a point should be considered.

Of course, there are no hard-and-fast rules for what such a range would look like. The ranges set forth in the Sentencing Table, for example, are simply a function of the "25% rule" created by Congress, and nothing more.[48] These ranges have no empirical or theoretical basis; they are a simple function of a simple mathematical calculation intended to confine the exercise of judicial discretion to a narrow range.

Given the significant amount of sentencing data now available to analyze for most cases, the Interquartile Range—the box in the box plots above—for the class of similarly situated individuals should be used by courts as the starting point for considering where to impose a particular sentence. As discussed above, the Interquartile Range ("IQR") is the difference between the first and third quartiles and thus encompasses where half of all similarly situated offenders have been sentenced. By definition, the IQR encompasses the median sentence and almost always the average sentence too. The IQR, unlike the Sentencing Table, is both descriptively and prescriptively accurate. It describes where courts are imposing sentences on similarly situated individuals, and prescribes where they should be imposed for the typical offender in order to achieve reasonable uniformity.

So, for example, and referring back to Figure 3, if a court is confronted with a defendant who is to be sentenced under USSG § 2B1.1, has a TOL of 21 in CHC I, and is not subject to a statutorily mandatory or consecutive sentence, then rather than looking to the sentencing range of thirty-seven to forty-six months set forth in the Sentencing Table as the starting point, a court should look to the IQR instead. In this case, it would be twenty-four to thirty-seven months.

Using the IQR as the starting point can address the three primary goals of the guidelines. First, because the IQR is necessarily based on sentences imposed on similarly situated individuals, a sentence within the IQR would achieve reasonable uniformity. Indeed, over time, it would help normalize sentencing and possibly reduce the spread of the IQR as more sentences are imposed within it. Second, it would better achieve proportionality than the Sentencing Table, given that most offenders are being sentenced below the Sentencing Table's range. More to the point, the IQR reflects where courts nationwide believe a sufficiently serious sentence lies. Finally, sentencing consistent with the IQR promotes certainty to the extent that it provides a better prediction of the ultimate sentence. Pre-*Booker*, once the guidelines TOLs and CHCs were calculated, the Sentencing Table generally gave the defendant, the attorneys, and the court a good idea of what sentence would be imposed—they just had to refer to the Sentencing Table. As we now know, the Sentencing Table is largely irrelevant for most offense types, especially as the TOL increases. Using the IQR as, in effect, a new Sentencing Table will provide at least a far better estimate for all parties involved as to the likely sentence.

VI. Conclusion: Harmonizing Sentencing through Statistics

Since *Booker*, over one million federal defendants have been sentenced under the guidelines. As the late Justice Scalia correctly predicted, a merely advisory guidelines system inevitably resulted in a wildly discordant symphony of sentences. But the data collected from these sentences reveal a method for harmonizing federal sentencing practice with the primary purposes of the guidelines.

By using the guidelines calculations to identify similarly situated defendants, the IQR for that class of offenders can then be determined. Sentencing consistent with the IQR will rationalize sentencing over time and necessarily promote the primary purposes of the guidelines: uniformity, proportionality, and certainty. Appellate courts would also benefit from this approach in determining whether particular sentences are substantively unreasonable. As at least one circuit court of appeals already has found when considering Commission statistics, even sentences imposed within a particular guidelines sentencing range can still be substantively unreasonable.[49] Likewise, significant downward variances also can be warranted in consideration of relevant statistics.[50] Should courts begin in earnest to actually impose heartland sentences on heartland offenders,[51] this would presumably provide the impetus to the Commission to recalibrate the Sentencing Table to bring it more in line with actual sentencing practice, rather than engage in piece-meal tinkering with individual guidelines. As far as sentencing practice is concerned, the data can (and should) act like a digital common law of sentencing, thus reducing the cacophony currently confronting the courts.

The guidelines can and should be saved. Similarly situated offenders ought to be transparently sentenced similarly, but proportionately to the seriousness of their offenses. The often overlooked, robust sentencing data that are collected and published by the Commission serve as a means—if not *the* means—for harmonizing the discordant sentencing symphony.

Notes

* The author is a former staff attorney to the U.S. Sentencing Commission. He can be reached at mark@sentencingstats.com.

1 *See* Brent E. Newton & Dawinder S. Sidhu, *The History of the Original United States Sentencing Commission*, 45 Hofstra L. Rev. 1167, 1175 (2017).

2 *See* Marvin Frankel, Criminal Sentences: Law without Order 113–14 (1972).

3 Pub. L. No. 98-473, 98 Stat. 1987 (codified as amended in scattered sections of Titles 18 and 28 of the U.S. Code).

4 *See generally* USSG Ch. 1, Pt. A(1) (original introduction to the Guidelines Manual).

5 *See* Mistretta v. United States, 488 U.S. 361, 367 (1989).

6 All references to "Sentencing Table" refer to the sentencing table set forth in Ch. 5, Pt. A of the U.S. Sentencing Guidelines Manual.

7 *See* USSG §§ 5K1.1, 5K3.1; Fed. R. Crim. Proc. 35.

8 *See* 28 U.S.C. § 994(o), (p).

9 543 U.S. 220 (2005).

[10] *Id.* at 244 (often referenced as the Constitutional Opinion).
[11] *Id.* at 233; *id.* at 245 (often referenced as the Remedial Opinion).
[12] *Id.* at 312 (Scalia, J., dissenting).
[13] *Id.* (quoting Remedial Opinion at 265).
[14] *See* USSG Ch. 1, Pt. A(3). Interesting, the guidelines never adopted a particular sentencing philosophy. *See* Paul J. Hofer and Mark H. Allenbaugh, *The Rules: Finding and Using the Philosophy of the Federal Sentencing Guidelines*, 40 Am. Crim. L. Rev. 19, 25 (Winter 2003).
[15] *See* USSG Ch. 1, Pt. A(4)(b).
[16] 543 U.S. 220 (2005).
[17] *Id.* at 233.
[18] *Id.* at 245–46.
[19] 542 U.S. 296 (2014).
[20] Variances, of course, did not exist prior to *Booker*. The pre- and post-*Booker* comparative periods are for only five years as the Commission changed how it reports such departures and variances in 2003 and 2018 in Table 28.
[21] Rosales-Mireles v. United States, 138 S.Ct. 1897, 1904 (2018).
[22] Molina-Martinez v. United States, 136 S.Ct. 1338, 1348 (2016).
[23] *See id.* at 1345.
[24] Peugh v. United States, 569 U.S. 530, 545 (2013).
[25] 569 U.S. 530 (2013).
[26] 482 U.S. 423 (1987).
[27] *Id.* at 540.
[28] United States v. Garrison, 560 F. Supp. 2d 83, 84 (D. Mass. 2008).
[29] *See* United States v. Williams, 788 F. Supp. 2d 847, 889 (N.D. Iowa 2011).
[30] *See* USSG §§ 5G1.1, 5G1.2.
[31] *See id.*
[32] *See* U.S. Sentencing Comm'n, Inter-district Differences in Federal Sentencing Practices: Sentencing Practices across Districts from 2005–2017, at 9–14 (Jan. 2020) (hereinafter "Inter-district Differences"), available at https://www.ussc.gov/research/research-reports/inter-district-differences-federal-sentencing-practices.
[33] The citizenship of an offender often can also be a critical factor to consider, especially when the applicable TOL is low enough that the guidelines provide for alternatives to imprisonment such as probation and community confinement. For example, offenders with TOLs below 12 (and thus in Zones A and B) are eligible for noncustodial sentences. However, an offender who is a deportable alien is generally not eligible for alternatives to imprisonment. *See, e.g.,* United States v. Martinez-Ramos, 184 F.3d 1055, 1057 (9th Cir. 1999) (status as deportable alien made defendant "ineligible for community or home confinement, and minimum security imprisonment"). As a result, the presence of deportable aliens can skew statistical findings upward in such circumstances. Because the analyses in this article are of offense levels well above 12, the citizenship of the offender is not considered.
[34] I chose to focus the analysis on those offenders who had received a TOL of 21 in CHC I inasmuch as this class of offenders was sufficiently large after applying the specified criteria to perform the analyses below.
[35] *See* Inter-district Differences at 10.
[36] *See id.* at 63, n.70.
[37] *See id.* at 63.
[38] *See id.* at 41 ("[T]he analyses in this report demonstrate increasing variations in sentencing practices across districts in the wake of the Supreme Court's 2005 decision in *Booker*.").
[39] United States v. Boucher, 937 F.3d 702, 713 (6th Cir. 2019) (quoting United States v. Stock, 685 F.3d 621, 629 n.6 (6th Cir. 2012)).
[40] United States v. Trejo, 729 Fed. Appx. 396, 401 (6th Cir. 2018) (citing United States v. Mullet, 822 F.3d 842, 854 (6th Cir. 2016)).
[41] *Mullet*, 822 F.3d at 854.
[42] The focus of § 3553(a)(6) is on "national disparities, not specific individual cases." United States v. Gamble, 709 F.3d 541, 555 (6th Cir. 2013) (citing United States v. Wallace, 597 F.3d 794, 803 (6th Cir. 2010)).
[43] To more accurately reflect the rate of the exercise of judicial discretion to vary below the applicable sentencing range, all sentences subject to a mandatory minimum or consecutive sentence were eliminated, as were all sentences that were the result of a government-sponsored motion for a downward departure.
[44] Excluding any offenders subject to a mandatory minimum or consecutive sentence, or who had received a government-sponsored motion for a downward departure or variance.
[45] The Commission records life sentences as 470 months.
[46] *See* USSG § 5G1.2(d).
[47] Again, excluding any offenders subject to a mandatory minimum or consecutive sentence, or who had received a government-sponsored motion for a downward departure or variance.
[48] *See* 28 U.S.C. § 944(b)(2).
[49] *See* United States v. Jenkins, 854 F.3d 181, 189 (2d Cir. 2017) (reversing a within-guidelines sentence as "shockingly high" based a review of Commission statistics).
[50] *See* United States v. Musgrave, 647 Fed. Appx. 529, 538 (6th Cir.) (affirming downward variance from 57–71 month sentencing range to one day imprisonment and two years' home confinement "[b]ased on the district court's review of statistics and other cases") (citing *inter alia* Mark H. Allenbaugh, *"Drawn from Nowhere": A Review of the U.S. Sentencing Commission's White-Collar Sentencing Guidelines and Loss Data*, 26 Fed. Sent'g Rep. 19, 19 (2013)).
[51] *See* USSG CH. 1, Pt. A(4)(b) ("The Commission intends the sentencing courts to treat each guideline as carving out a 'heartland,' a set of typical cases embodying the conduct that each guideline describes.").