Amoral Numbers and Narcotics Sentencing

Mark Osler
AMORAL NUMBERS AND NARCOTICS
SENTENCING

Mark Osler*

I. INTRODUCTION

We Americans are strangely drawn to numbers. We pore over football rankings before any team has played a game. We endlessly discuss “Ten Best” lists of songs or movies or television shows, ignoring the sheer subjectivity of that entire process. In my own field, we treat the largely artificial law school rankings as if they came down from an objective (and judging) God.

In each instance, the numerical orders we obsess over pretend to describe an existing hierarchy. In truth, they are creating one. A team becomes the “preseason favorite” and that shapes further outcomes. A dress is listed as a “hot new look” on the cover of a magazine and therefore becomes one. A law school rises in the rankings for reasons having nothing to do with quality and that undeserved image of quality becomes their identity, leading to the enrollment of more and better qualified students—at least for a year.

This attraction to numerical systems, which create rather than describe reality, too often leads us to false priorities and bad decisions. Such is the case with the numerical matrices at the center of federal criminal law: the arbitrary mandatory minimums and sentencing guidelines that rank-order the severity of crimes have time and again created broad and often tragic outcomes in our society. This dysfunction is perhaps most clearly seen in the laws and guidelines governing narcotics. These drug sentencing rankings too often are unrooted in anything of substance,¹ yet have created a new and troubling reality—the destruction of communities, the imprisonment of thousands of citizens, and the spending of billions of dollars, all of which create the baseless illusion that we are “doing something” about illegal narcotics.²

¹ For example, the Supreme Court described the sentencing guidelines for crack cocaine as resulting from a process that “did not take account of ‘empirical data and national experience.’” See Kimbrough v. United States, 552 U.S. 85, 109 (2007) (McConnell, J., concurring) (quoting United States v. Pruitt, 502 F.3d 1154, 1171 (10th Cir. 2007)).

² See generally Jessica M. Eaglin, Neorehabilitation and Indiana’s Sentencing Reform Dilemma, 47 VAL. U. L. REV. 867 (2013) (exploring the incarceration boom in Indiana and proposing a new model of criminal justice reform); Brian G. Gilmore & Reginald Dwayne Betts, Deconstructing Carmona: The U.S. War on Drugs and Black Men as Non-Citizens, 47 VAL. U. L. REV. 777 (detailing the ways in which the war on drugs has affected African-American communities).
This Article seeks to describe the problem of these criminal law systems, which create rather than reflect reality, and to suggest a better way that will acknowledge our fascination with numbers yet turn us away at least from the worst outcomes of our current and wildly deceptive system of normative rankings, such as sentencing guidelines and mandatory minimum sentences.

Part II broadly describes how our society is drawn to numerical systems that are often unmoored from an existing reality, yet often create one, and provides three examples. Most strikingly to those of us in legal education, the rankings of law schools are one good example of this effect.

Part III moves to specifics and establishes how it is that the numerical sets at the center of criminal law steer us wrong. This examination centers on four types of numerical measures: sentencing guidelines, mandatory minimum sentences, the weight of narcotics possessed as a mistaken proxy for culpability in drug trafficking, and the number of narcotics cases as well as the resulting sentences as a flawed indicia of success in drug interdiction.

In turn, Part IV suggests a better way to use numbers in the field of narcotics interdiction, one that will move towards solving the problem of narcotics use rather than simply towards mass incarceration. Numbers are not the enemy, but neither have they proven to be the solution. As we re-examine the way in which we address narcotics, part of that evaluation should focus on the careless way we have let numerical systems become not just tools, but a shimmering hologram of principle itself, directing action even though there is no humanity within.

II. THE ALLURE OF NUMBERS

Before looking at how mandatory minimums and sentencing guidelines in criminal law create rather than reflect a hierarchy of seriousness, it is appropriate to examine how this works in other areas. Criminal law, it turns out, is not alone in suffering from this phenomenon.

I’m not the first to note the preponderance of meaningless rankings; as usual, the satirists got there first. My current Senator, Al Franken, deftly made his point about the way a veneer of objectivity can attach to even the most subjective judgments once they are translated into a

---

ranking or list by publishing his own “World Religions in Order of Quality.” It is a joke, of course (it’s a comedy book, after all), but the religious quality rankings play out intriguingly:

1. Judaism (Reform)
2. Judaism (Conservative)
3. Unitarianism
4. Christianity (Mainstream Protestant)
5. Christianity (Roman Catholic)
6. Islam (Muhammad Ali/Ahmad Rashad type)
7. Buddhism, Hinduism, Confucianism, etc.
8. Judaism (Orthodox)
9. Christianity (Fundamentalist)
10. Islam (Fundamentalist)

Having looked over this list, I ask you to engage in a simple reflection. Despite the obvious subjectivity, possible offensiveness, and even the silliness of this list, announced beforehand, didn’t you find yourself either agreeing or re-arranging the list mentally? In other words, didn’t you, reading this simple, obvious hoax, engage with it (even if just for a moment) as if it was real? That is the power of lists. There is something within us that longs for order, even if that order is baseless. We can take random things, compile them in a list, and that ordering somehow becomes authentic.

Nor can we avoid lists and ranking. They are everywhere. Fashion magazines are full of numbered lists of what is “in,” and those lists do more to create popularity than to reflect it. This, unfortunately and dangerously, extends beyond fashion itself to the body types within those fashions. In the realm of college football, dynasties are built in part on the strength of preseason rankings—supposedly descriptive lists which instead create future realities through their effects on recruits and program resources. Finally, and closest to home, the U.S. News and World Report rankings of law schools does much more to create images of quality than it does to reflect the actual and distinct strengths of existing schools with enormous collateral damage to the operation of law schools in the United States.

4 AL FRANKEN, OH, THE THINGS I KNOW!: A GUIDE TO SUCCESS, OR, FAILING THAT, HAPPINESS 50 (2002). The book also includes a ranking of “places to hide once the shooting starts” and an exhaustive list of “things you love” indexed to “what it causes.” Id. at 121, 139. The latter list contains the claim that watching The Bold and the Beautiful causes both obesity and diabetes. Id. at 139.
A. Marketing and Magazines

No one knows what appeals to us more than marketers (that is their business, after all), and they frequently use lists, completely baseless lists, to draw us in. For example, the designers of InStyle Magazine, a magazine that is often prominently displayed near supermarket checkout counters, carefully include the promise of some kind of list on the front cover of every issue. Most issues contain more than one; for example, the May 2012 cover for InStyle offers up not only a picture of Cameron Diaz, but headlines for “12 Beauty Secrets Nobody Tells You,” “92 Cute Spring Outfit Ideas,” and “187 Style Finds Under $50.”

This profusion of lists isn’t accidental; there are no accidents on the “front door” of a carefully tended magazine. Marketers know that numbers draw us in and lend a sheen of authenticity to nearly anything—even something as inherently subjective as an evaluation of “cute spring outfits.”

When a fashion magazine sets out a list of “92 Cute Spring Outfits,” of course, it is also proactively defining, in the public eye, those particular spring outfits as being “cute.” It is hard to imagine a more subjective measure than “cuteness.” In fact, just about the only public measure of what kind of spring outfit is “cute” is just such a list. The objectivity comes from the precise and somewhat odd number of cute spring outfits: exactly ninety-two. At first glance, this may seem utterly harmless.

But only at first glance. This normative function has a dark side. Those cute outfits too often look cute only on a certain kind of body, and therein lies the rub. Fashion magazines have come under attack for their role in creating unrealistic body images in women, and especially in teenage girls. In May 2012, one teenage girl, Julia Bluhm, staged a protest outside of the offices of Seventeen magazine, asserting that the magazine airbrushed photographs and that pictures in the magazine “did not represent real adolescent females, and contributed to unattainable ideals[,]” a charge that resonated with the public.

---

Numbers on a magazine cover, then, help to create a new reality, and it is that new reality which can be directive to society as a whole. The same is true in other areas as well.

B. Football Rankings

Every August, before a single college football game is played, a slew of preseason rankings are issued. Most prominent among these are the listings compiled by the Associated Press and USA Today. The polls manage to include what turn out to be some spectacular flame-outs among teams ranked highly in the preseason, even when other parts of the poll are strikingly accurate. For example, in 2012, the Associated Press poll accurately predicted the regular season finish of three preseason top ten teams: Alabama at number two, Oregon at number five, and Georgia at number six. It also chose to promote, as the top team in the nation, a group which then went on to lose six games, including a loss in a third-tier bowl played in El Paso.

Of course, all the teams ranked in the preseason poll, even the flame-outs, were, in a sense, big winners. That’s because rankings not only reflect the quality of a school’s team, they play a role in creating and maintaining the quality of the team. Highly ranked teams play more often on television, which both generates more money and entices recruits. Many newspapers commonly list and report only on games

---

8 The AP and USA Today rankings are used as part of the BCS rankings, which determine major bowl eligibility, though the BCS poll itself does not issue preseason rankings. The 2012 preseason rankings are available at Randy Chambers, College Football Preseason Rankings 2012: AP and USA Today Top 25 Preview, BLEACHER REPORT (Aug. 27, 2012), http://bleacherreport.com/articles/1312409-college-football-preseason-rankings-2012-ap-and-usa-today-top-25-preview.

9 See id. (listing the University of Southern California as number one in the Associated Press’s poll). For example, in 2012, the preseason number one team, the University of Southern California, did not finish in the top twenty-five at the end of the regular season. 2012 NCAA Football Rankings – Postseason, ESPN, http://espn.go.com/college-football/rankings (last visited Jan. 24, 2013).


11 See John Erford, Georgia Tech Beats Southern California in Sun Bowl, ASSOCIATED PRESS (Dec. 31, 2012, 11:51 PM), http://collegefootball.ap.org/article/georgia-tech-beats-southern-california-sun-bowl-0 (explaining that the University of Southern California football team, ranked number one in the preseason, lost in the Sun Bowl to a Georgia Tech team that was only able to play in a bowl because of a special waiver from the NCAA, required because of Georgia Tech’s losing regular-season record).
involving local teams and those in the top twenty-five—meaning that you have to be in the top twenty-five to be reported on nationally.

Importantly, all this is true not only of end-of-the-year rankings, but also of the preseason rankings, which are made without the benefit of any games having been played and involve teams with tremendous turnover in personnel from year to year. Thus, even the seemingly meaningless preseason rankings create a new reality—they determine who goes into the season with confidence, who gets on television, who is covered nationally, and who is most attractive to recruits.

To understand this effect, consider the biggest anomaly of the 2012 season, the University of Southern California—the team that ended with that loss in El Paso. They started the season ranked first in the nation but ended up with a (relatively) mediocre 7-5 regular-season record, which included losses to Arizona, Notre Dame, Oregon, Stanford, and UCLA. Despite this seeming disaster, no one doubts that USC will be a powerful team next year and into the future. One reason for this is that USC got verbal commitments from top national high school players early on in the 2012 season, while they were still riding the wave of what turned out to be an undeserved national ranking. In other words, USC’s reality in future years will in part be created by the completely inaccurate and speculative preseason rankings in 2012. The ranking created a reality (good recruits and likely future success) even as it proved so dismal at reflecting the reality of USC’s strength.

Just as a completely arbitrary list of “cute summer outfits” can play a role in body image problems, so these arbitrary football rankings create problems—albeit, not as serious as the health issues related to bulimia and anorexia. Because the preseason rankings tend to list the same schools every year, they play a role in the continuation of a rigid hierarchy within the sport. Even as a down year, for example, USC got a boost from the preseason rankings, and the highly-ranked recruits they scooped up early in the season will not be distributed to lesser teams, depriving them of upward mobility. There are, after all, only a limited number of players in the United States who are both big and fast. These lesser teams continue to spend outrageous amounts of money to play

---

14 In the same way, the teams that were accurately predicted to succeed in 2012 by the preseason poll (Alabama, Oregon, and Georgia, for example) will also be strengthened for the future.
big-time football, yet those schools are fighting a structure that constantly tilts against them regardless of merit.

C. Law School Rankings

In the spring of each year, U.S. News and World Report issues its annual ranking of law schools. It creates this listing by blending a number of factors. Most important are two surveys of reputation. A full quarter of the ranking is based on a “peer assessment” survey sent to some law professors, while another 15% flows from the results of a similar survey sent to lawyers and judges. The remainder of the score is accounted for by assessments of job placement rates nine months after graduation (14%), the LSAT scores of incoming students (12.5%), expenditures per student (11.25%), undergraduate GPA’s of incoming students (10%), job placement at graduation (4%), student-faculty ratio (3%), the acceptance rate of the school (2.5%), bar passage rate (2%), and library resources (.75%).

Virtually every aspect of this survey has come under attack as either unreliable or subject to manipulation by law schools. The “peer review” portions, for example, which make up a large percentage of the ranking criteria, are remarkably “sticky”—that is, they seem to stay the same from year to year regardless of what is actually happening at a given school. Moreover, the peer review by practicing lawyers and judges suffers from a terribly low response rate: meaning that in an average year fewer than 200 lawyers and judges have determined this important part of the rankings. Shockingly, this means there are about the same number of law schools in the United States as there are voters.

16 Id.
17 Id.
18 One of the more persuasive critiques has come from Washington University Professor Brian Z. Tamanaha, whose book, Failing Law Schools, contains a myriad of acute observations relating to the U.S. News rankings. BRIAN Z. TAMANAHA, FAILING LAW SCHOOLS (2012).
19 This tendency extends even back before the advent of the U.S. News Rankings. Brian Leiter of the University of Chicago has catalogued the static nature of reputational rankings back to 1977 on his blog. Brian Leiter & Dan Filler, The More Things Change, the More They Stay the Same . . . Again, BRIAN LEITER’S LAW SCHL. REPORTS (Mar. 13, 2012), http://leiterlawschool.typepad.com/leiter/2012/03/the-more-things-change-the-more-they-stay-the-sameagain.html.
20 TAMANAHA, supra note 18, at 79–80. U.S. News has been forced to begin averaging over two years in this category. Id.
in a sector of the rankings that provides a full 15% of the final tally.\textsuperscript{21} The other categories are also flawed, largely in that they invite “gaming” by the schools to increase their rank. Law schools, for example, have included any kind of employment and even hired their own graduates to inflate “employment” figures,\textsuperscript{22} and have adjusted their policies on transfers to maintain revenue without jeopardizing the LSAT numbers that count for the U.S. News rankings.\textsuperscript{23} That is nothing, though, compared to the outright lying about numbers engaged in over a period of years by the law schools at Villanova and the University of Illinois.\textsuperscript{24}

The pressure to do well has twisted the priorities of law schools in some gruesome ways. Instead of using financial aid to help needy students, grants are strategically employed by many schools to harvest an incoming class with relatively high LSAT numbers and grades.\textsuperscript{25} The result at schools which do not have huge endowments is that the students most likely to get high-paying employment are subsidized by the law school debt accumulated by their lower-scoring classmates.\textsuperscript{26} That is, if scholarships are funded by tuition money, it is the less-qualified students who subsidize those with better credentials. The shocking meta-story here, combining the relative meaninglessness of the rankings’ inputs and the importance of the results, is that the ranking does far more to create outcomes than it does to reflect or report them.

For example, Professor Tamanaha points to the experience of Emory Law School, which dropped eight places from the 2011 rankings to those released in 2012.\textsuperscript{27} This drop did not correlate to any significant change at that school, but it represented a tremendous driver of outcomes. The dean resigned.\textsuperscript{28} Perhaps more importantly, it put Emory in the position of having to put far more resources into luring the very best students in terms of LSAT scores (rather than the neediest ones) just to maintain the hope of returning to its former status, because the eight-point drop was sure to produce a drop in applications and the qualifications of

\textsuperscript{21} \textit{Id.} The U.S. News rankings released in 2012 listed 200 schools, and this did not include those schools that are not accredited. \textit{Best Law Schools}, U.S.\textsc{news}, http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-law-schools/law-rankings/page+8 (last visited Jan. 24, 2013).

\textsuperscript{22} TAMANAH\textsc{A}, supra note 18, at 71–72.

\textsuperscript{23} Id. at 88–94.

\textsuperscript{24} Villanova and Illinois falsified data on the qualifications of incoming students. \textit{Id.} at 74–76. See also Richard W. Painter, \textit{Numerical Half Truths, Human Lies, and Other Distortions of Truth}, 47 \textsc{Val. U. L. Rev.} 479 (2013) (contending that the scandals at the University of Illinois and Villanova were, at least in part, a result of “our obsession with the numbers”).

\textsuperscript{25} TAMANAH\textsc{A}, supra note 18, at 96–99.

\textsuperscript{26} Id. at 98.

\textsuperscript{27} Id. at 80.

\textsuperscript{28} Id.
applicants. In turn, this furthers the unfortunate focus on LSAT scores, rather than the life stories, financial need, or actual abilities of applicants.

As with fashion magazines and football rankings, law school rankings create a new reality, which then takes on a life of its own, built on smoke and mirrors.

III. NORMATIVE RANKINGS AND LISTS IN CRIMINAL LAW

In the fields described above, we saw how lists and rankings create interest, direct outcomes, and give meaning to things even when there is little or no empirical basis for the formulation of those lists themselves. The same is true in criminal law.

The three examples given above (fashion magazines, college football, and law schools) all involve lists that relate in some way to marketing of a product or service. Fashion magazines are selling not only the magazines, but the styles promoted by advertisers. College football is selling tickets and seeks television revenue, all of which are benefits from increased exposure. For law schools, U.S. News rankings are often used in marketing materials, and the rankings themselves sustain a formerly robust news organization, which has largely been reduced to a ranking service.

But what about criminal law? Certainly, rankings such as those found in the Federal Sentencing Guidelines are not driven by the same kind of consumerist marketing, since there is nothing being bought and sold (at least not directly). However, there is a common impulse at work. Sentencing guidelines and mandatory minimums are both the creation of elected officials, people who are often anxious to be seen as doing something about a particular type of crime. Just as a marketer who needs to sell a dress would be well advised to have it be listed in InStyle Magazine as a “cute summer look,” or a dean looking to increase the status of his law school might try to game U.S. News ranking numbers, so a politician who wants to seem thoughtful and forceful about crime might look to create that image through the seeming objectivity of numerical matrices. “Five years for five grams of crack” may not really be that different than “92 Cute Spring Outfits.”

Here, we will examine four types of lists or ranking devices within criminal narcotics law at the federal level: sentencing guidelines, mandatory minimum sentences, the use of weight as a proxy for culpability, and the use of the number of people incarcerated or drugs seized as a measure of success. Taken together, these numbers-based systems have a tremendous influence on the operation of criminal law.

29 Id.
Indeed, as caseloads grow, it is easy to understand how numerical systems would seem like a good idea, as they tend both to promise uniformity in sentencing from case to case and greater efficiency.

I focus here on federal systems not because they are most important, but because they are capable of being described within the scope of this Article. State experiences vary widely, and others have already done an excellent job of cataloguing this diversity. Moreover, the federal experience has defined anti-drug efforts in the national media, in part because narcotics captured the attention of Congress. As Frank Zimring put it, “The lead role in declaring a war on drugs was played by Congress, which has throughout the period after 1914 played a much larger role in penal policy regarding drugs than any other forms of crime.”

If we are looking for a national view of sentencing trends and effects, there really is only one national system of sentencing.

A. U.S. Sentencing Guidelines

1. Hierarchy Within the Guidelines

The U.S. Sentencing Guidelines were implemented in 1987. For nearly two decades they were mandatory, but in 2005 the Supreme Court ruled in United States v. Booker that to preserve the constitutionality of the guideline system they would be “advisory” to sentencing judges. Despite this advisory status, the Sentencing Guidelines continue to have great weight in federal sentencing. The structure of the sentencing process still largely centers on a narrow sentencing range produced by these guidelines, and there are still substantial disincentives to varying from the guideline range.

One of the more fascinating features of these guidelines is the formal ranking it necessarily creates, putting in order the perceived relative severity of various crimes. The Guidelines are based around a grid that lists sentencing ranges according to two inputs: the seriousness of the offense, which defines a numerical “offense level,” and the defendant’s

---

30 Franklin E. Zimring, Sanctions: Penal Policy and Penal Legislation in Recent American Experience, 58 STAN. L. REV. 323, 331–332 (2005). Zimring notes that more than a quarter of the people imprisoned nationally for drug crimes were federal prisoners, compared to about four percent of those imprisoned for “common violent and property crimes.” Id. at 332 (footnote omitted).


33 See Mark Osler, Seeking Justice Below the Guidelines: Sentencing as an Expression of Natural Law, 8 GEO. J. L. & PUB. POL’y 167, 172–73 (2010) (discussing several disincentives to varying from the Sentencing Guidelines). These disincentives include a desire to avoid reversal, the public tracking of sentences, and the risk of inviting greater restrictions. Id.
prior criminal history. Offense levels are calculated starting from a “base level,” to which are added various points for aggravating factors (or, more rarely, points are subtracted for mitigating factors). For example, the base offense level for a simple kidnapping is thirty-two points. To that base level, six points are added if a ransom demand is made, four points are added if the victim suffers permanent injury, and six more points are added if the victim is sexually exploited, among other factors.

In contrast, insider trading is a less serious crime under the Federal Guidelines with a base offense level of eight, with enhancements available based on the amount of money the defendant gained from the scheme. The widely varying offense levels between these two crimes produce very different sentences under the Guidelines. For example, if we look at the simplest form of each (that is, without enhancements), as committed by someone with no criminal history, the result is that the trader would be subjected to an advisory guideline range of zero to sixty months, while the kidnapper would face a sentence of 121–151 months.

At a very basic level, then, the Sentencing Guidelines create a hierarchy, which is revealed through relative offense level scores. At the top of that hierarchy will be the most serious crimes and at the bottom will be the least serious crimes, because we naturally want to give the harshest punishments to those who commit the most serious crimes. The Guidelines themselves refer to this ordering of seriousness as “proportionality.” To order crimes into a hierarchy of seriousness and then call that ordering “proportionality” certainly gives the result an air of scientific precision. This impression is reinforced by the sheer heft of the U.S. Sentencing Guidelines Manual (now expanded to three thick volumes) and the impressive progression of punishment set out in the sentencing grid. However, that air is nothing more than a vapor, especially when we look at the Guidelines’s sense of proportionality as applied to narcotics sentences.

The story behind the creation of this “proportional” hierarchy was revealed by one of its framers, now-Justice Stephen Breyer of the U.S.
Supreme Court, who was a member of the U.S. Sentencing Commission from 1985 to 1989 while serving as a judge on the First Circuit Court of Appeals. Writing in the Hofstra Law Review about the process the Commission used to draft the first set of guidelines, Justice Breyer laid bare a completely unscientific and in fact quite messy process. It appears that almost immediately there was conflict on the Commission as “different Commissioners [had] different views about the correct rank order of the seriousness of different crimes.... [T]he members of the group inherently tend[ed] to ‘trade’ over particular items so that each person [found] his own views reflected only some, but not all, of the time.” From Breyer’s report, it seems that the commissioners were reluctant to “abandon their own subjective values” and also expressly rejected the systemic use of a more objective ordering mechanism, such as public polling or an assessment of the possibility of deterrence.

Deadlocked on this issue, Breyer reveals that “the Commission reached an important compromise”: rather than trying to objectively order the ranking of offense levels, they simply used data from previous sentencings to mold the new guidelines from past practice, and then adjusted them where they saw fit. There was simply no objective principle at work in the end.

2. Narcotics Within an Unprincipled Guideline Hierarchy

As described above, the Commission that framed up the Federal Sentencing Guidelines did not use any meaningful sorting principle to establish the hierarchy of seriousness that defines the offense-conduct input to those Guidelines, other than an amalgamation of past practices. This, in turn, left a blank slate (in a moral sense) for later commissions and Congress, which had the power to alter that hierarchy but were left without consistent or even articulable guidance on the basis

45 Id. at 15.
46 Id. at 16.
47 Id. at 17–18.
48 In critiquing this process, I do not mean to denigrate any use of sentencing systems that rely on collecting broad sentencing data, then making that available to judges or others, provided that this process is used consistently as time goes on, rather than as the baseline for a system which will then steadily increase that baseline independent of systemic input from sentencing judges. In fact, I have advocated for real-time sentencing information systems to be used in place of sentencing guidelines such as that in use in federal courts. Such real-time systems would likely have the effect of moderating harsh sentences. Mark Osler, The Promise of Trailing-Edge Sentencing Guidelines to Resolve the Conflict Between Uniformity and Judicial Discretion, 14 N.C. J.L. & TECH. 203 (2013).
for that structure or its amendment. Flooding into that vacuum, reactive politics filled the void. Not surprisingly, Congress proceeded to create harsh drug sentences both in statutes and the Guidelines in a frighteningly herky-jerky style, largely reacting to public reports about new drug “epidemics.” The result was a guideline book that boasted a stringent, yet baseless, sense of proportionality where narcotics sentences regularly exceeded the terms of nearly every other type of crime.

The bizarre outcome of this process is easily observed. In the Guidelines of today, for example, the base offense level for distributing 300 grams (about ten ounces) of crack cocaine or 500 grams of methamphetamine is thirty-two, the same offense level we saw for kidnapping. While this amount is far more than anyone would need for personal use, it is hardly remarkable; large narcotics rings in a large city might go through several times this amount in a single day. Yet, the ranking afforded this offense, as reflected in the offense level of thirty-two, is greater than that given for the forcible rape of an adult (thirty), killing a person in a voluntary manslaughter (twenty-nine), disclosing top secret national defense information to the North Koreans (twenty-nine), arson creating a substantial risk of death (twenty-four), and extortion (eighteen).

As I have set out elsewhere, one fascinating aspect of this dysfunctional ranking system is that it often punishes more harshly the narcotics crime that indirectly victimizes people (through increased violence, for example) than it does those direct victimizations. In other words, one reason to punish drug activity is because it leads to things

---

49 Primarily through mandatory minimums, such as those included in 21 U.S.C. § 841.
50 Congress at times has directed changes to the Sentencing Guidelines themselves and in 1995 vetoed a proposed change that would have lessened sentences for crack cocaine. U.S. SENTENCING COMM’N, SPECIAL REPORT TO THE CONGRESS: COCAINE AND FEDERAL SENTENCING POLICY 2 (1997), available at http://www.ussc.gov/Legislative_and_Public_Affairs/Congressional_Testimony_and_Reports/Drug_Topics/19970429_RTIC_Cocaine_Sentencing_Policy.PDF.
51 U.S. SENTENCING GUIDELINES MANUAL § 2D1.1 (2012).
52 Justice Breyer’s Hofstra article implies that the Sentencing Commission in the end decided not to create rankings relating to the severity of crimes. See Breyer, supra note 44. However, that ignores reality; regardless of how the Commission thought of its work, the bare fact that plays out in courtrooms daily is that the offense levels assigned to various crimes are constructed as a numerical matrix. It is this ordinal ranking that plays a large role in determining the sentences that people actually serve.
53 U.S. SENTENCING GUIDELINES MANUAL § 2A3.1(a)(2).
54 Id. § 2A1.3(a).
55 Id. § 2M3.3.
56 Id. § 2B3.2(a).
like random violence (manslaughter) and rape, yet we punish narcotics trafficking more harshly than we do the actual commission of manslaughter and rape under the hierarchy established by the Guidelines. The drug trafficker described above has an offense level of thirty-two—two more than the rapist, and three more than the person who commits manslaughter. Drug dealing, it seems, is not only seen to be an indirect cause of rape and killing, but somehow worse than actually raping and killing (at least in some circumstances).

3. How Guidelines Create Reality

The reality surrounding someone who sells 300 grams of crack is not very remarkable. A low-level street dealer who sells ten grams every day hits that standard in a single month after all. There is nothing about that particular street-level dealer that distinguishes him from thousands of others or makes him any more dangerous than other drug dealers. In fact, it is hard to imagine anyone beyond the most casual and temporary distributor not meeting this threshold if all relevant conduct is considered. Yet, the Guidelines create a remarkable status for him—serious offender worthy of very harsh punishment relative to others—that is unrelated to any objective standard. This status simply does not correlate to facts on the ground, given the probability that this particular crack dealer is necessarily buying powder cocaine from someone else, who is buying it from a network of others, who likely bought it from an importer. Viewed holistically, the only thing that would define this bottom-feeder as a serious offender relative to those others is the artifice of the Sentencing Guidelines and the related numerical systems, such as mandatory minimums.

In other words, this unsubstantiated status of serious offender doesn’t exist until it is created by the Guidelines.

However, once that status is created by the Guidelines, it has very real consequences. The defendant is much more likely to get a long sentence, and that sentence is less likely to be overturned on appeal. We taxpayers get to fund the housing of this defendant for that lengthy term. One consequence we won’t see is this: any difference in actual drug trafficking. Because street-level trafficking has such low barriers to entry, the position will quickly be filled and nothing will change other than a slight bump in American incarceration.

B. Statutory Mandatory Minimum Sentences

Statutory minimum sentences are closely intertwined with the Sentencing Guidelines in the federal system; for example, a mandatory minimum will establish the effective floor for a guideline where both apply except in certain defined circumstances. However, these minimums are distinct from the Guidelines in at least two important ways. First, they comprise a less comprehensive ranking system than the Guidelines because they cover far fewer crimes. Second, they are even more directive than the Guidelines because they are mandatory in most cases.

Like the Guidelines, the statutory minimums set out in the federal code hit hard at even minor narcotics offenders. The most-often used mandatory minimum sentences are contained in 21 U.S.C. § 841, which provides for stiff sentences for distributing a variety of narcotics. There are two primary groups subjected to mandatory minimums under 21 U.S.C. § 841(b), which Congress referred to as “serious” traffickers and “major” traffickers. For example, possessing with the intent to distribute 280 grams of crack cocaine makes one a “major” trafficker and earns you a minimum sentence of ten years (in the absence of prior narcotics convictions—if you have one of those, the minimum doubles to twenty years), while just five grams of methamphetamine gets a five year minimum sentence (or ten years with a prior conviction) as a “serious” trafficker.

One odd feature of the mandatory minimums that perhaps pops out more clearly in the statute than in the Guidelines is the often arbitrary way in which drugs are valued in relation to one another. For example, a five-year minimum applies to both ten grams of PCP and five grams of methamphetamine, despite the fact that both drugs have the same amount per dose, five micrograms. Why is it that methamphetamine is

---

59 U.S. SENTENCING GUIDELINES MANUAL § 5C1.2. See generally Lynn Adelman, The Adverse Impact of Truth-in-Sentencing on Wisconsin’s Efforts to Deal with Low-Level Drug Offenders, 47 VAL. U. L. REV. 689, 689 (2013) (“I have long been convinced that the federal sentencing laws and guidelines result in an enormous amount of over-punishment, particularly in drug cases.”).


62 Id. § 841(b)(1)(B)(viii).

63 Id. § 841(b)(1)(B)(vi) (PCP).

64 Id. § 841(b)(1)(B)(vii) (Methamphetamine).

65 U.S. SENTENCING GUIDELINES MANUAL § 2D1.1 cmt. 9 (2012).
taken twice as seriously as the very dangerous PCP? And why are both of them ranked so far ahead of heroin, for which it takes a whopping 100 grams to meet the five-year threshold of this same provision?

There really is no good answer. The ranking of methamphetamine being more serious per dose than PCP, which in turn is more serious than heroin, is abjectly arbitrary. Yet, the numbers, drawn from vapor, create a hierarchy that has real effects and drives the engine of over-incarceration.

One aspect of mandatory minimums that extends beyond the effects of the guideline matrix is the way that not only narcotics but the possession of firearms in relation to narcotics trafficking is subjected to harsh sentences, even when the firearm isn’t used, or even shown. One of the sharpest denunciations of these minimums came in an award-winning opinion by Judge Paul Cassell of Utah. In that case, United States v. Angelos, a first-time narcotics offender ran into a welter of mandatory guidelines after being convicted of two small-time ($350) marijuana sales and the possession of three guns (which were not used or brandished in the course of the crimes). Because of these statutory minimums, Judge Cassell was compelled to impose a fifty-five-year sentence.

In his opinion, Judge Cassell not only expressed his disagreement with the arbitrariness of the statutory minimums (calling them “unjust, disproportionate to his offense, demeaning to victims of actual criminal violence—but nonetheless constitutional”), he called for a presidential commutation of the sentence he was giving. Moreover, Judge Cassell did something remarkable, unusual, and illuminating: he critiqued the mandatory minimums by bringing to the surface the false ranking of seriousness that they contain. Specifically, Judge Cassell included in the

66 PCP is so dangerous that, according to the National Institute on Drug Abuse, clinical trials on humans could not be completed because of negative effects on the subjects. See PCP/Phencyclidine, NAT’L INST. ON DRUG ABUSE (last updated Dec. 2012), http://www.drugabuse.gov/drugs-abuse/pcpphenacyclidine.
70 Angelos, 345 F. Supp. 2d at 1227.
71 Id. at 1230.
72 Id. at 1261.
73 Id. at 1262–63.
opinion itself two tables which compared the sentence required in the relatively minor Angelos case to the treatment of other crimes under the guidelines and mandatory minimums, noting that the defendant’s marijuana trafficking received a much higher mandated sentence than crimes such as causing death as part of narcotics trafficking, hijacking an aircraft multiple times, and raping ten-year-old children.74

A sitting judge is at the very point where the unprincipled numbers meet the realities of a human life, and Judge Cassell’s opinion is compelling in large part because it exposes the raw and human effects of that interaction.

C. The Use of Weight as a Ranking of Culpability

Built into the mandatory minimums and Sentencing Guidelines described above is an equally unsupportable and much older ranking system—the assignment of sentences based on the weight of the narcotics a defendant may possess with the intent to sell. For example, a federal defendant who holds just twenty grams of cocaine and is sentenced pursuant to the relevant guideline range faces an offense level of twelve,75 while one who is caught with sixty kilograms is rated at an offense level thirty-six.76 This translates, for defendants with no criminal history or other relevant factors, into a guideline sentence of ten to sixteen months for the former offender and 188–235 months for the latter.77 This enormous gap exists because of the completely unfounded assumption that the person who possesses the most narcotics is necessarily the most culpable. This idea is undone by the simple reality of mules: people who ferry large amounts of drugs from one place to another for relatively low pay. They possess a lot of narcotics but are not very culpable in the grand scheme of things. Moreover, incarcerating mules, like incarcerating street dealers, will make almost no difference in drug trafficking, because they are easily replaced by other low-wage laborers.

By ranking the culpability of narcotics felons primarily based on the weight of the narcotics they possess or possessed, we are using a very crude measure of culpability and creating incentives to go after people who really don’t matter much. The Sentencing Commission’s own data reveals that this incentivization is exactly what is happening, with the result that federal resources are too often being used primarily to pursue people who are street dealers, mid-level managers, or mules. For

74 Id. at 1247.
75 U.S. SENTENCING GUIDELINES MANUAL § 2D1.1(c)(14) (2012).
76 Id § 2D1.1(c)(2).
77 Id. at ch. 5, pt. A., tbl.
example, the Commission’s report to Congress on crack cocaine crimes in 2007 revealed that 66.9% of crack defendants in federal court were prosecuted for these lower-level actions, which combine low culpability (by any realistic measure) with weight amounts significant enough to trigger higher sentences. Like the Guidelines and mandatory minimums generally, the use of weight as a proxy for culpability within those systems does not reflect a reality; rather, it creates one that serves no real social purpose.

D. Numbers of Convictions/Seizure of Narcotics

Press releases in federal drug cases consistently trumpet three things: the amount of narcotics seized, the number of people arrested or convicted, and the lengths of the sentences they received. Big numbers in any of these areas—big piles of marijuana, twenty-seven defendant indictments, 360-month sentences—impress us. We are accustomed to thinking in terms of numbers (as already established) and these kinds of numbers often appeal to us visually, as well. We respond to the image of a stack of marijuana bales or cocaine bricks, for example, with a certain sense of satisfaction; it means that something is being done about the drug problem.

Of course, this is often patently untrue. We assume that the measure of success is amount of marijuana seized, or people arrested, or length of sentence, but if what we care about is actually solving the drug problem, none of those numbers are particularly useful because the drugs and the people are often easily replaced. As it turns out, the sense of success these big numbers convey is an illusion. Seizing stacks of marijuana and cocaine, bringing in lots of arrests, and getting long sentences haven’t resulted in success against the problem of narcotics trafficking and use. The federal government’s own analysis shows that the use of narcotics in this country remains at high levels, virtually unchanged from previous years.


80 See Pot King Gets 27 Years in $6M Bust; McChesney Faced a Minimum of 10 Years in Prison, WANE.COM (Mar. 4, 2011, 3:47 PM), http://www.wane.com/dpp/news/matthew-mchesney-sentencing-marijuana-pot (providing images of drugs that were confiscated during a bust in Fort Wayne, Indiana).
The drug war, thus far, has been won by drugs, and the kind of enforcement we have pursued may actually increase violence rather than reduce it.82

We have assumed that if we seize a lot of drugs or imprison a group of people that those drugs or people won’t be immediately replaced, but too often there is no grounding in reality for those assumptions. Once again, numbers, even in the absence of substance, create a reality: the impression that something worthwhile is being accomplished, which serves to garner continuing support for the institutions and people that create this illusion.83

IV. AN ARGUMENT FOR BETTER NUMBERS

A. A Better Number with Which to Evaluate Interdiction: Price

As the examples in Part II above show, informally compiled lists and rankings perform three functions. They mask subjectivity with a veneer of credibility. They draw interest. Perhaps most important, they create a new reality and a set of outcomes that flow from that reality rather than what actually pre-existed the list or ranking. We see all of these things at work in American criminal law within the federal system. In the end, we are left with an odd combination: numbers that seemingly announce a remarkable string of successes and a tough approach on crime, and little to no change in the actual problem being addressed.

This Article is not meant as a general screed against the use of numerical evaluation—just the use of numerical systems that are made up without a firm basis in reality. After all, without numerical

83 The focus on numbers in federal sentencing, and in particular the way the sentencing process centers on calculating a guideline range from a number of numerical inputs, also serves to greatly reduce the role allowed to a traditional tool of the oppressed: narrative. While defendants are generally afforded a chance to speak at sentencing, this story-telling moment is often secondary or tertiary and mooted by the number-driven process. The relating of individualized circumstances is controlled and directed by the numerical matrix of the Sentencing Guidelines, rather than the historical sorting mechanism for story-telling, which is the passage of time. By reducing individualized circumstances to particularized numerical inputs, federal sentencing takes leave of a rich and worthwhile ancient tradition in sentencing, which allows the full story of the defendant, told holistically and chronologically, to play a central role in the broad consideration of an appropriate punishment.
evaluations, such as stock prices, our economy would grind to a halt. What is wrong with the rankings described here is that they don’t correlate to real markets, and it could be that the solution lies with a better use of economics and market analysis.

After all, the nonsense numbers described in Part II (“92 Cute Spring Outfits”) would hardly withstand an economic analysis, because the inputs are so purely subjective or are unrelated to any current reality or guiding principle for which there might be a societal consensus. Instead, the numbers create the illusion of a societal consensus in order to create an effect. We are seduced by them, and seduction has little to do with rational thinking and logic.

So what might an economic evaluation tell us about narcotics? At a very basic level, an economic analysis would focus on markets, and the measure of a market is price. Absent sudden changes in demand for a product, if you increase the supply of that product, the price will go down. If you decrease the supply, the price will go up. Given that, the true measure of success in drug interdiction (which aims to diminish supply) isn’t how many people you lock up, it is the price of that product in a market.

Thus, the real evaluation tool for gauging our narcotics interdiction efforts is street price. If that is true (and I think it is), the news isn’t good. The price of cocaine and heroin actually went down during the height of the “war on drugs.”

Thus, the reality created by the primary normative number systems described here—sentencing guidelines and mandatory minimums—was not one that negatively impacted narcotics trafficking by the best economic measure, price.

B. Using a Number that Works

As set out above, numbers of people incarcerated and pounds of drugs seized are a lousy gauge of success in drug interdiction, and mandatory minimums and tough sentencing guidelines don’t do much to increase the price of drugs. If that is the case, what should be done to eradicate drugs?

---

Just as economics and a focus on markets offer us a better method of evaluation, it offers us a better means of interdiction, as well. The narcotics trade, after all, is a business, and businesses are governed by the rules of economics.

To start with, consider an accessible and worthwhile chapter from perhaps the most popular economic book of the last few decades: 2005’s *Freakonomics*, by Steven D. Levitt and Stephen J. Dubner, devoted a fascinating chapter to an analysis of the crack trade, titled, “Why Do Drug Dealers Still Live with Their Moms?”85 There, based on information from the inside of a drug-selling gang in Chicago, they determined that street-level crack dealers in Chicago were making only about $3.30 an hour.86 However, it was (like many other jobs in that gang) a role that nearly anyone could do without special resources or talents. The business, as a whole, has what economists call “low barriers to entry.” The effect of low barriers to entry within the business of narcotics at the bottom levels means quite simply that sweeping up low-wage labor will not make much difference. That, unfortunately, has been at the forefront of our anti-drug efforts: sweeping up low-wage labor in the person of mules, street dealers, and mid-level managers. This effort has been sustained by the cynical manipulation of meaningless numbers as described above—by claiming that the possessor of five grams of methamphetamine is a “serious” trafficker, that the Guidelines really express some rational ranking of harm, and that weight is a true proxy for culpability.

A better method to raise the price of drugs (if that is what we want), one rooted in business practices rather than a moral crusade, would be to focus on the cash flow returning to bulk producers of narcotics. In other words, ignore the people and take the money. Unlike the people we imprison, that cash flow is not so easily replaced because drug dealers do not have ready access to legitimate credit.87 At the very least, this plan would offer a less destructive focus for our obsession with rankings and numerical evaluation.

86 Id. at 103.
IV. CONCLUSION

We can’t, and shouldn’t, completely avoid the use of numerical matrices and rankings in our lives. However, we can be careful about how such numerical systems are used, and we have done a lousy job of this in the realm of criminal law. Rather than carefully creating something like the Sentencing Guidelines or mandatory minimums based on a well-thought-out ordering system that is rooted in reality, we have created them thoughtlessly and without principle. This allows the rankings contained in guidelines and mandatory minimums to direct action and justify outcomes, despite the fact that they are rooted in nothing.

The cost of this mistake is profound and extends beyond the brutal cost of over-sentencing. The price of fake rankings also includes the loss of any comprehensive moral basis for this important social function. Without principles and reality informing these most directive and dramatic societal actions, the law is nothing but a bully.