Law & Society REVIEW

Legal Change and Sentencing Norms in the Wake of *Booker*: The Impact of Time and Place on Drug Trafficking Cases in Federal Court

Mona Lynch Marisa Omori

The federal sentencing guidelines have lost some authoritative force since the U.S. Supreme Court ruled in a series of recent cases that the guidelines are advisory, rather than presumptive, in determining criminal sentences. While these court decisions represent a dramatic legal intervention, sociolegal scholarship suggests that organizational norms are likely to change slowly and less dramatically than the formal law itself. The research reported here looks specifically at the consequences of such legal transformations over time and across locale, using multilevel analysis of U.S. Sentencing Commission sentence outcome data from 1993 to 2009. Our findings suggest that districts vary considerably from each other in sentencing practices over the time period studied, and that there is relative within-district stability of outcomes within districts over time, including in response to the Supreme Court's mandates. We also find that policy change appears to influence the mechanisms by which cases are adjudicated in order to reach normative outcomes. Finally, we find that the relative district-level reliance upon mandatory minimums, which were not directly impacted by the guidelines changes, is an important factor in how drug trafficking cases are adjudicated. We conclude that local legal practices not only diverge in important ways across place, but also become entrenched over time such that top-down legal reform is largely reappropriated and absorbed into locally established practices.

Introduction

The authoritative force of the federal Sentencing Guidelines has been diminished by a series of legal decisions in recent years. Most notably, the U.S. Supreme Court in the consolidated cases of *U.S. v. Booker and U.S. v. Fan-Fan* (2005), then *Gall v. U.S.* (2007), *Kimbrough v. U.S.* (2007), and *Pepper v. U.S.* (2011) ruled that the Guidelines are no longer mandatory, but merely advisory in determining

We would like to thank Amy Baron-Evans, John Hipp, Richard McCleary, Lou Reedt, Elina Treyger, as well as the LSR editors and anonymous reviewers for their thoughts, comments, and advice. An earlier version of this article was presented at the Conference on Empirical Legal Studies, Stanford Law School. This project was supported by Grant No. 2010-IJ-CX-0010 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Please direct all correspondence to Mona Lynch, UC Irvine, Criminology, Law and Society, 2311 Social Ecology 2, Irvine, CA 92697-7080, USA; e-mail: lynchm@uci.edu.

Law & Society Review, Volume 48, Number 2 (2014)

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

criminal sentences in federal court. The cumulative effect of these decisions is that judges may now consider a broad range of sentencing factors in determining an appropriate sentence, including those not specified in the Guidelines or in conflict with its policy statements. Thus, judges are free to impose sentences other than those prescribed, so long as they both begin with the Guidelines calculations, and then explicitly state their reasons for any variance from the Guidelines' range.

These rulings have the potential to dramatically reshape federal sentencing, as judges had been relatively hamstrung from exercising individualized sentencing discretion since the Guidelines were put into effect in 1987. Yet previous sociolegal scholarship suggests that norms about appropriate case outcomes are likely to change slowly and less dramatically than the formal law itself, while perhaps the language and mechanisms for negotiating to those normative outcomes might change shape. Moreover, the way that federal district courts adapt to this transformed legal landscape is likely to vary by locale, so the decisions' cumulative impact would thereby diverge in intensity, quality, mechanisms, and outcomes. In this article, we examine whether and in what ways these formal legal changes have reshaped legal practice in federal courts.

We conceptualize the research as comprised of two key dimensions—time and place. Our analyses track sentencing over time in an effort to measure how the *Booker* line of cases, along with other sentencing policy changes, affect sentencing behavior on the ground. Integrated into our conceptual model is the notion that place also matters—both in terms of local legal structures and norms, as well as broader sociopolitical cultural norms. We view the federal district court system not as a singular national legal structure with hierarchically arranged and geographically dispersed subunits, but rather as a semi-autonomous set of systems governed by the same formal rules, statutes, and procedural policies, while also embedded in localized legal cultures that are themselves shaped by regionally specific historical contingencies and norms (Church 1985; Eisenstein & Jacob 1977; Ulmer 2005). Thus, we begin with the supposition that sentencing practices can potentially change not only over time in a pattern that reflects national-level impacts, but that districts and regions may reflect qualitatively and quantitatively different reactions to macro-level forces, like the *Booker* et al. legal mandates.

We limit our analyses to U.S. Sentencing Commission (USSC) federal drug trafficking case outcome data from fiscal years 1993–2009. The starting point of the time period was constrained by data quality and availability; nonetheless, this represents among the most comprehensive longitudinal regression studies done of the

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

USSC sentencing outcome data to date. We limited our analysis to drug trafficking for several reasons. First, since the inception of the Guidelines up until 2009, drug trafficking was the single largest category of sentenced cases in the federal system, so it comprises a significant share of federal criminal court interest and resources. In addition, drug trafficking sentences grew dramatically more punitive under the Guidelines, more so than any other category of offense; and federal drug policy, particularly in relation to crack cocaine, has been controversial both within the court community as well as among the general public (Lynch 2012).

From a conceptual standpoint, there is also reason to expect that the way in which drug trafficking offenses are adjudicated will differ from how other categories of federal offenses (such as immigration offenses and white collar economic crime) are treated, so our approach provides a cleaner test of change. As such, we heed Engen's (2011) suggestion that research examining the effect of *Booker/Fan-Fan* et al. on sentencing should consider the specific impact of the cases as a function of offense type. Finally, and importantly, drug trafficking cases are especially likely to be subject to a range of mandatory minimum statutes that have been enacted by Congress since the 1980s and that remain in full force, even after the Guidelines lost their presumptive status. Drug traffickers in particular are subject to the two sets of sentencing regimes—the Guidelines, which increasingly allow for individualized and nuanced sentencing, and mandatory minimum statutes that rely upon very limited offense criteria (such as weight of drug involved) to mandate a sentence outcome.

The Federal Sentencing Guidelines and the Underlying Ideals for Sentencing Regulation

The United States Sentencing Commission (USSC) was established in 1984 with the passage of the Sentencing Reform Act, and one of its primary purposes was to develop a set of sentencing guidelines that would help decrease sentencing disparities between similar cases and across jurisdictions, and ensure consistency and transparency throughout the system. The new sentencing structure was supposed to strictly limit the range of possible outcomes for "like" defendants, increase "certainty" that defendants would be punished, and increase the severity of penalties for certain offense categories (USSC 2009). Within 3 years of this mandate, the Commission had drafted an intricate and rigid set of guidelines that was put into effect on November 1, 1987. The Guidelines have since undergone numerous changes and additions, which are reflected in each annual Guidelines Manual.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

The Guidelines have also been supplemented by a number of statutory mandatory minimums and mandatory enhancements that have been passed by Congress on a regular basis since 1984. When applicable, these statutory provisions are binding and must be taken into account when determining sentence; judges may not sentence below them even if they would like to do so. The Sentencing Commission chose to incorporate mandatory minimums into Guidelines' sentencing grid, by increasing Guidelines ranges across offense levels to meet the statutory dictates, and to eliminate "cliffs" in which sentences change dramatically by offense level (see Baron-Evans and Stith [2012] on how mandatory minimums were incorporated into the Guidelines grid).

The Guidelines Manual, which provides the specific rules, policies, and procedures for calculating sentencing ranges for federal criminal offenses, currently runs over 500 pages (with indexes, appendices, supplements, and tables, it runs over 2,000 pages) and requires intensive training for those who do the calculations for the court, primarily U.S. probation officers. The formulae are for the most part additive, where offense levels easily go up through "relevant conduct" assessments and "specific offense" characteristics and enhancements, resulting in dramatic increases in federal imprisonment rates and sentence lengths since the 1980s (Stith & Cabranes 1998). Conversely, "acceptance of responsibility" and playing a minor role in the offense are among the very few factors that can lower offense levels.

The Guidelines and related federal sentencing statutes are illustrative of a larger phenomenon in criminal justice administration that took place in the late twentieth century whereby individual offenders' motivations, deficiencies, needs, and potential have become irrelevant to the sentencing decision (Feeley & Simon 1992). Instead, the focus shifted to assessments of aggregated offender classes, and efficiency and consistency in administration. This trend is epitomized by sentencing schemes that use criminal history as a key determinate of sentencing, and those that constrain judicial discretion in the sentence decision (Feeley & Simon 1992). While the stated purpose of sentencing under the Guidelines is a mixed bag, including "just punishment, deterrence, incapacitation, and rehabilitation" (USSC 2009: 1), the Commission aimed to largely take the unique defendant—aside from criminal history—out of the process, thereby eschewing an individualized approach to sentencing. Indeed, as Savelsberg (1992) has argued, the Guidelines represent a retraction from substantive rationality in the criminal law to formal rationality, particularly through its elevation of uniformity (based on offense characteristics and criminal history) over other justice goals.

The Guidelines constrained individualized sentencing through a number of design features that reduced judicial discretion, and that devalued the individual defendant in the formulation. First, the determination of penalties is derived from a 258-cell grid that captures two distinct aspects of the offense and defendant. The vertical axis of the grid is the "offense level," which ranges from 1 to 43 and is calculated by considering statutory conviction, specific offense characteristics (such as weight of drugs involved in a trafficking offense, whether reflected in the conviction or not), any statutory mandatory minimum or enhancement "trumps," specified aggravating and mitigating factors, and so on. The horizontal axis is a "criminal history score" ranging from 1 to 6, which is derived from a formula that assigns points to prior criminal convictions, temporal relation to most recent custodial release, and current criminal justice supervision status. This structure was the Commission's solution to the alleged problem of irrationality and uncontrolled discretion in the system, ideally resulting in consistency and uniformity in sentencing based on an "empirical approach" to sentencing (USSC 2010a: 4).

Second, the Guidelines deemed most aspects of the defendant's background (other than criminal history), present circumstances (except the circumstances of the offense, even if unproven in court), and future needs or potential as irrelevant to the sentencing decision. The Guidelines Manual prohibits, as a matter of policy, consideration of a whole host of traditionally relevant defendant-specific sentencing factors for purposes of "departure," including age, educational attainment or vocational skills, employment history or career potential, family status or responsibilities, physical, mental or emotional condition, and disadvantaged background. The sentencing grid offense levels for a given conviction also generally ignore criminal intent, and instead impose a strict liability standard of culpability. Thus, the amount of drugs, money, or other illicit goods involved in the larger offense largely drives sentence severity even if the accused had direct involvement in just part of the larger criminal offense.

Sentencing in Practice

Inherent in the Guidelines structure is the assumption that sentencing is controllable by a set of rules that will successfully homogenize interpretations of complex events and contingencies. In other words, the Manual's dictates are expected to have the same effect across diverse cases, decision makers, and places. But the presumption that legal policy is translated into practice in a uniform and orthodox manner is belied by a long line of law and

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

society scholarship. Indeed, the question of whether there is a gap between "law on the books" and "law in action" has long ago been asked and answered (Abel 2010); more fruitful questions interrogate the specificities of translating formal law into practice. In the context of criminal courts, now-classic scholarship has demonstrated the impact of proximate and distal contexts on how cases are adjudicated (see Ulmer 2012 for a full review). For example, local courts and the workgroup members within them, share enduring norms about the terms of, and routes to, adjudication (Church 1985; Eisenstein & Jacob 1977). Adjudication norms in criminal courts are typically formed as to "going rates" for prototypical cases, what cases are triable, how long a case should take until resolution, and the terms of negotiations in plea deals (Church 1985; Dixon 1995; Eisenstein et al. 1988; Heumann 1978; Kramer & Ulmer 2002; Ulmer & Johnson 2004). The specifics of these norms are conditioned and constrained by the structure and mandates of the larger jurisdiction's formal law, but practices vary considerably by intra-jurisdictional locale (Johnson et al. 2008; Lynch 2011).

Broader sociopolitical arrangements and historical intergroup relations that transcend jurisdictional lines can also imbue local criminal justice operations, impacting how cases are resolved. For example, Southern U.S. locales have traditionally exhibited more punitiveness in both formal policies and in actual practices compared with other regions, while the Northeast has been relatively more lenient and rehabilitative in orientation (Lynch 2011). There are also norms about criminal offense types that eclipse jurisdictional borders. Thus, sentencing leniency persists for certain white collar defendants, relative to those convicted of "street crime," even in the wake of the highly publicized white collar crime scandals of the 2000s (Maddan et al. 2012; Van Slyke & Bales 2012).

More fundamentally, a long line of sociolegal research has highlighted how frontline criminal justice actors, including police officers (Grattet & Jenness 2008), prosecutors (McCoy 1993; Nagel & Schulhofer 1992; Schulhofer & Nagel 1989, 1996), and penal staff (Kemshall & Maguire 2001; Lynch 1998), implement legal policy change in ways that can dramatically reshape and even subvert formal goals and directives. Indeed, in bureaucratic institutions like those that comprise the criminal system, frontline workers are often isomorphic agents, actively resisting top-down reform mandates that would alter established operating logics and procedures (Cheliotis 2006; Robinson 2002), so new rules get adapted and absorbed into established practices (Feeley & Kamin 1996).

Federal courts are not immune to the influence of such factors, even in the most restrictive periods of mandatory sentencing

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

schemes. Rather, those actors with the power to put law into action will reshape it in ways that serve their institutional purposes and reinforce their legitimacy. In this case, despite the intentions of the Guidelines, the sentencing reforms catalyzed a shift in discretionary power, from judges to U.S. attorneys, making its exercise less transparent and more adversarial given the prosecutor role in the courtroom workgroup configuration. As Berman (2010: 429) suggests, prosecutors became "first-look sentencers" in the context of the mandatory sentencing schemes, as charging decisions came to largely dictate sentence outcomes. And because prosecutors can control much of the information about the offense, they have some power to circumvent, or drive up, the "relevant conduct" provision of sentencing through the information they release to probation officers who calculate the Guidelines ranges (Shermer & Johnson 2010; Wilmot & Spohn 2004).

Moreover, in cases subject to mandatory minimums, federal prosecutors hold the sole power to authorize sentences below the minimum. The primary vehicle for that authorization is through a 5 K1.1 motion by the prosecutor recommending a sentence below the minimum because the defendant provided "substantial assistance" to the government. No formalized rules govern the degree of reduction, nor define what qualifies as substantial assistance, so these values are subject to variation across sentenced defendants, prosecutors, districts, and time (Ulmer 2005). Because this mechanism rewards benefits according to how much information is given and the prosecutors' valuation of it, rather than as a matter of relative culpability, it can contribute to significant disparities between otherwise similar cases (Ulmer 2005; Wu & Spohn 2010).[1] The other route to relief is the application of the "safety valve," which releases low-level drug defendants who have no more than 1 criminal history point from statutory mandatory minimums. In practice, this can function as another government-controlled substantial assistance mechanism, as it requires eligible defendants to fully "debrief" about the offense, including about codefendants and others, in order to obtain a prosecutor's recommendation for a reduction. Thus, it is also subject to variation in use and relative value across actors and locales.

While much of the empirical scholarship on federal sentencing has focused on whether individual-level outcomes became more

[1] On the issue of justice and culpability, the USSC's (2011) most recent report on mandatory minimum sentencing indicates that in drug trafficking cases, street-level dealers had the lowest chance of all traffickers to get relief from the mandatory minimum for any reason, including by providing substantial assistance, whereas high-level managers, wholesalers, and importers/manufacturers were comparatively likely to get relief overall, and especially through providing substantial assistance to the government.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

uniform under the Guidelines (with a particular interest in whether "extralegal" demographic factors continued to shape sentences; e.g., Albonetti 1997; Doerner & DeMuth 2010; Mustard 2001), several studies have examined the sociolegal contexts of districts. Schulhofer and Nagel conducted two early, mixed-methods studies assessing whether and how the Guidelines were being circumvented through the plea bargaining process (Nagel & Schulhofer 1992; Schulhofer & Nagel 1989, 1996), finding that circumvention varied significantly in frequency and in type between districts (Schulhofer & Nagel 1996). In a more recent cross-sectional study, Kautt (2002) documented significant interdistrict variability, finding that the organizational context alone, as represented by district caseload factors and circuit-level appeal rates, could not fully account for the wide variation observed. Rather, she found that local legal cultural factors better explained observed geographic variations in sentencing.

Johnson et al. (2008) also examined interdistrict variability in federal sentencing processes, focusing on the frequency and degree of prosecutor-initiated substantial assistance departures and judge-initiated downward departures. They found evidence that both organizational characteristics, as well as broader sociopolitical characteristics of districts were associated with patterns of use for both types of departures. Finally, Wu and Spohn (2010: 298) tested the "uniformity assumption" across three contiguous midwestern districts, which they characterized as a "conservative" test of the assumption. Nonetheless, they found significant differences between the districts, particularly in the type, frequency, and value of departures, thereby contributing to differential sentence outcomes.

Ulmer's (2005) mixed-methods research provides an especially rich picture of how districts vary in sentencing processes and outcomes under the presumptive Guidelines. Using qualitative interviews with courtroom workgroup members coupled with quantitative analyses of outcome data for four distinct districts, he documented processual differences between the districts, which differentially impacted case outcomes. Thus, while outcome data indicated significant differences between the districts, the interview data revealed just how much the underlying definitions of the terms of adjudication vary by local setting. Indeed, the qualitative data pinpointed where and how outcome data alone mask some fundamental differences by locale. Ulmer (2005: 272) concluded that despite the intent of the Guidelines to provide uniformity across disparate districts, "court community actors interpret Guidelines and other federal criminal justice policies differently, and use and transform these in a variety of ways."

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

The Guidelines Become Guidelines

As previously noted, several recent Supreme Court decisions ruled that the presumptive federal Guidelines scheme violates the Sixth Amendment of the Constitution. The remedy ordered in *Booker/Fan-Fan* (2005) was to excise the statutory language that made the Guidelines mandatory, thereby rendering them "effectively advisory" (245). Despite this ruling, the USSC, and some courts, continued to treat the Guidelines as presumptive, and even nearly fully mandatory. In response, the Court reiterated its position that the Guidelines are merely advisory, when it ruled in *Kimbrough v. United States* (2007) that judges are free to sentence outside of the prescribed Guidelines' range on the grounds of policy disagreements with the Guidelines. In the contemporaneous case of *Gall v. United States* (2007), the Court imposed a standard of review mandating deference to sentencing judges' decisions, effectively freeing judges to use individualized assessments of cases and defendants in deciding whether and how to depart from the Guidelines (see Baron-Evans and Stith [2012] for a full discussion).

An emerging body of empirical scholarship has looked at the impact of the *Booker* line of cases on sentencing outcomes. The USSC itself has published three reports on the impact of *Booker.* In the year immediately following *Booker*, the Commission's analysis indicated that district courts' sentencing practices had not changed dramatically from the immediate pre-*Booker* period (Hofer 2007; USSC 2006). In its 2010 analysis, however, the USSC inferred that there has been an increase in sentence length disparity between similar defendants, particularly as a function of race (USSC 2010b).[2] In its 2012 follow-up on the "continuing impact of *United States v. Booker*" (USSC 2012: 3), the Commission reported that "unwarranted disparities in federal sentencing appear to be increasing."[3]

Studies conducted by scholars who are not associated with the Commission suggest a more equivocal state of affairs. For instance, Jeffery Ulmer and his colleagues (Ulmer et al. 2011a,b) conducted alternative analyses of sentencing change brought on by *Booker* et al., replicating the 2010 USSC's study design, but including additional relevant controls in their model; disaggregating the prison sentence length from the prison/no prison sentence

[2] While scholars have called into question the methodological choices made in this study (e.g., Ulmer et al. 2011a), this analysis prompted the Commission to develop policy reforms aimed at reining in the judicial discretion afforded by *Booker* et al. (see Baron-Evans & Stith 2012; Saris 2011).

[3] See Hofer (2013), a former senior research associate at the Commission, for a pointed critique of this study.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

decision; and lengthening the period under study to more fully capture the mandatory Guidelines era. Generally, their findings indicated that racial disparities were equal or less than those observed in previous periods, with the one exception that some nonwhites' *rate* of incarceration increased after *Gall* relative to Whites (Ulmer et al. 2011a,b). Moreover, they found that "unwarranted" disparities in the post-*Booker* period continue to be, in large part, a product of prosecutorial behavior, and that any changes brought about by the legal rulings "defy easy characterization into a convenient narrative" (Ulmer & Light 2011: 339; see also, Hofer 2007, 2011).

Rehavi and Starr (2012) conducted a set of analyses designed to address the shortcomings of looking for "unwarranted" disparities only at the final stage of the criminal justice process, as a large share of variation can result from charging decisions and early-stage plea negotiations. The researchers used data from multiple stages of process to examine how *Booker* impacted both charging and sentencing, finding that observed post-*Booker* racial disparities are largely due to prosecutorial behavior. Specifically, directly following *Booker,* prosecutors were more likely to lodge charges subject to mandatory minimums, but only against black defendants, thereby increasing the charging stage racial disparity. Over time, prosecutors began to file more mandatory minimum charges against whites. They did find that judicial departures increased after *Booker*, but equally for blacks and whites.

Fischman and Schanzenbach (2012) more directly tested the emergence of observed racial disparities in sentence outcomes over time, also finding that increased sentence disparity between blacks and whites is largely due to prosecutors' use of mandatory minimum statutes. Specifically, disparities emerge as judges bump up against those minimums in sentencing, especially in the post-*Booker* period. Their findings "suggest that judicial discretion does not contribute to, and may in fact mitigate, racial disparities in Guidelines sentencing" (Fischman & Schanzenbach 2012: 761). Finally, in terms of between-district variations, including whether *Booker* has increased racial disparities, Farrell and Ward's (2011) work suggests that analyses attempting to isolate the impact of formal legal change are complicated by other co-occurring changes to the federal system. In particular, the demographic diversity of federal courtroom workgroups decreased post-*Booke*r, which, as they have shown, impacts rates of racial disparity in sentence outcomes (Ward et al. 2009).

Both Rehavi and Starr (2012) and Fischman and Schanzenbach (2012) zero in on an important contextual consideration as to how cases are settled in the post-*Booker* era, which is that the mandatory minimums enacted by Congress in the 1980s and 1990s are left

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

untouched by the legal decisions. In cases subject to mandatory minimum sentences, the influence of the *Booker* line of cases is limited to the instance when the Guidelines sentence is lengthier than the statutory minimum, and its extent of influence stops at that minimum. In cases where the mandatory minimum is higher than the Guidelines sentence, *Booker* et al. should have no influence at all. Yet our understanding of how cases are resolved in trial-level courts would predict that when one constitutive element of sentencing is changed (i.e., the Guidelines), all elements of sentencing will be affected, and the rules of bargaining will holistically change such that mandatory minimum cases could inhere new meaning. Thus, it may be that U.S. attorneys elect to seek mandatory minimums in more eligible cases in the post-*Booker* era as a way to mitigate the uncertainties of judicial sentencing under the new "advisory" Guidelines system (Rehavi & Starr 2012). Additionally, "going rates" for mandatory minimum cases may well be reduced if increased judicial power to sentence outside of the Guidelines range drives down sentences in Guidelines cases.

As we delineate later, this study is designed to further unpack these complexities of the federal sentencing system as it undergoes policy transformation. We examine whether, and in what ways, the two sentencing tracks—advisory Guidelines and mandatory minimums—which are both applicable in the majority of drug trafficking cases, interact as legal mandates change over time. We also seek to assess, more generally, how districts sentence over time in response to changes to the formal rules of sentencing. Our conceptualization of sentencing in this project is derived from Ulmer's (2012: 8) characterization of sentencing events as: "joint acts produced by the discretion and interactions of judges, prosecutors, defense attorneys, and sometimes probation officers [that are] . . . embedded in (and maintain or change) local court communities, which are in turn embedded in local socio-cultural contexts." Therefore, we are not focused solely on measuring individual-level disparities in outcomes; rather, our questions aimed to look at sentencing more holistically to examine the translation processes of policy to practice by local court communities over time.

The Study

Methods and Data

We analyzed USSC sentence outcome data, supplemented with district- and state-level variables, to assess sentence outcome variation nationally, by district, for drug trafficking offenses subject to section 2D1 of the Guidelines, for the years 1993–2009. We ask a set

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

of research questions, delineated later, about the quality, variation, and extent of sentencing change over time:

1. To what degree do within-district sentencing patterns demonstrate stability across different policy periods, indicating the influence of local court norms?
2. How much do organizational factors explain district-level variance in sentencing? Do districts that handle large numbers of drug trafficking cases demonstrate less variation over time compared with districts that handle relatively fewer drug cases? Do districts with relatively heavier criminal caseload pressure demonstrate less variation across all time periods than those districts with lighter caseload pressure? Do districts characterized by high proportions of mandatory minimum drug cases differ from those districts with low percentages of mandatory minimum cases?
3. Have federal sentences in drug trafficking cases increasingly varied from the Guidelines sentence recommendations in the post-*Booker* periods? If so, are the increases significantly greater for cases that are not subject to mandatory minimums than for those that are subject to them? Or are sentence outcomes relatively stable across time, but the mechanisms for getting to those outcomes changing in response to policy reforms?

We obtained the Defendants Sentenced Under the Sentencing Reform Act data from the USSC from years 1993–2009. These data include information on individual defendants, including demographics, offense and criminal history categories, case characteristics, and final sentence outcome. Drug trafficking cases were identified using the sentencing guideline offense, which provided us with a sample of N = 280,954. Because we were most interested in variation in federal districts, we modeled the cases nested in the 89 federal districts.[4] We merged these data with federal district-level indicators from the Federal Court Management Statistics website, and supplemented them with state level demographic data from the U.S. Census Bureau, crime data from the Uniform Crime Report, and treatment admissions data from Treatment Episode Database.

We examined three measures of adjudication behavior, described later, to get a fuller picture of how sentencing works in practice over time, across policies, and as a function of district, and to minimize the risks associated with comparing outcomes only,

[4] We excluded Guam, the Mariana Islands, the United States Virgin Islands, Puerto Rico, and the District of Columbia. While we would have preferred to use the subunit of division (of which there can be three or more per district) for our level 2 unit of analysis, the USSC does not code by division.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

under different sentencing structures. As Piehl and Bushway (2007) have empirically demonstrated, outcome data under highly structured presumptive sentencing systems—like under the pre-*Booker* Guidelines—reflect less disparity than actually exists because the "charge bargaining" processes, which are especially likely to be used, occur fully outside of the formal court. They advise caution in comparing disparities among different kinds of sentencing schemes because researchers should "expect to find less measured disparity in studies of highly structured systems with conviction data than in more loosely structured systems even if both systems contain similar amounts of total disparity" (Piehl & Bushway 2007: 122; see also Savelsberg 1992).

We first examined mean final prison *sentence length* in months, transformed by taking the natural log to reduce skew.[5] The mean sentence length for drug trafficking across the cases in our sample was just over 80 months. Because we were interested in observing how sentencing may have changed with respect to adherence to the Guidelines, we also modeled the *percentage sentence difference* between the guideline minimum sentence and the actual sentence. The Guideline minimum accounts for the conviction and all other "specific offense characteristics"; criminal history; enhancements for weapons, priors, and other aggravators; and minimal role and acceptance of responsibility mitigators, as calculated in the presentence report. Because the difference between the Guideline minimum sentence and actual sentence had a very large number of 0 values (indicating that the actual sentences were at the Guideline minimum calculated by probation) but ranged from −470 to 470, we recoded this difference as a percentage of the actual sentence to the guideline minimum sentence. Thus, a value of 100 meant that the guideline minimum sentence and the actual sentence were the same (i.e., the actual sentence was 100 percent of the guideline minimum), values less than 100 indicated a smaller actual sentence compared with the guideline minimum sentence, and values greater than 100 represented defendants sentenced for longer periods of time than the guideline minimum sentence. We also captured relative variance from the Guideline minimum in this manner because using a raw difference score distorts the substantive value of a given deviation. Put simply, a 10-month departure is a much more significant reduction when the Guideline minimum is 30 months rather than 130 months. On average, drug-trafficking

[5] Per USSC's method, we recoded life sentences to 470 months. Additionally, we excluded probation, which was coded as "0" for prison time. Drug trafficking cases that resulted in probation alone were less than 5 percent of all sentences.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

defendants were sentenced to approximately 85 percent of the guideline minimum sentence length across the entire time period.[6]

Finally, we looked at *mandatory minimum* eligibility as a dichotomous outcome. Eligible cases included those in which, after conviction, the defendant was exposed to a mandatory minimum sentence. At that stage, as noted earlier, there are essentially two modes for relief from the mandatory minimum sentence. First, drug defendants with little or no criminal history who did not use weapons or violence may be granted the "safety valve," which reduces the offense level by 2 points and releases the defendant from the statutory minimum. The 2-point reduction is calculated into the final guideline minimum by pretrial probation, so does not measure as a deviation from presumptive sentence, but any further reduction below that would be captured as a deviation. The other mechanism for relief are government-sponsored motions, which are primarily comprised of 5 K1.1 motions for substantial assistance. These are granted after the Guidelines minimum has been calculated, so are measurable as a deviation from the Guidelines minimum. Importantly, 68 percent of drug trafficking cases faced mandatory minimum sentences at the sentencing stage and nearly one-third of those cases received substantial assistance reductions.

For our "time" questions, our independent variables of interest are the two time periods following the *Booker* case, on January 11, 2005, and the *Gall* and *Kimbrough* cases, on December 9, 2007. We also included time periods for pre-*Koon* (before June 1996), *Koon* (from June 1996–February 2003), and the PROTECT Act (March 2003–December 2004), as each of these periods represent distinct policy mandates related to federal sentencing. The pre-*Koon* period was characterized by increasing restrictions on judicial sentencing discretion, some of which was restored as a result of *Koon v. United States* (1996), establishing a more deferential standard of appellate review in departure cases. Congress responded to the subsequent perceived rise in percentage of sentencing departures by statutorily revising the standard of review in an amendment to the 2003 PROTECT Act (see Ulmer et al. 2011a, for a fuller discussion of this periodization). Consequently, the PROTECT Act period was the most restrictive on sentencing judges in the Guidelines era. The time period variables were

[6] The percentage of sentence difference is fairly skewed with many low values. We ran two alternative models, one by natural logging the outcome, and the other by running a count (negative binomial regression) model. However, the results were similar to the original models, so we present these results for ease of interpretation. We do want to raise caution regarding violating the assumption of normal distribution of errors, however.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

coded as dummy variables with the PROTECT period excluded as a reference category.[7]

We also included a number of individual case variables. In the sentence length and mandatory minimum models, we included recommended Guidelines sentence, measured as the guideline minimum. In the sentence length and sentence difference models, we included whether a mandatory minimum applied, as well as safety valve application. In all three sets of models, we included the defendant's criminal history category,[8] as well as the application of substantial assistance and other downward departures (coded as dummy variables). About 27 percent of drug trafficking cases in the total sample had substantial assistance departures, and 8 percent had other downward departures. In theory, some cases could have both substantial assistance and other types of departures, although this occurred in very few cases. Because sentences vary substantially by drug, we controlled for primary drug of crack cocaine, powder cocaine, heroin, marijuana, methamphetamine, and other drugs. These were modeled as a series of dummy variables, with crack cocaine excluded as a reference category. Finally, we included limited defendants' demographic information: gender, race/ethnicity (coded as dummy variables with white as the reference category), citizenship status, age, whether the defendant was a high school graduate, and number of children.

We expected districts to respond differently based on local caseload and organizational factors so we included these measures in our models. The criminal caseload size, adapted from Ulmer et al. (2010), is measured by the number (in hundreds) of criminal case filings per judgeship, in a year, and the district's relative efficiency was captured by the median time to disposition in months. To see whether and how case composition impacted sentencing, we included percentage of trafficking cases out of overall caseload. We aggregated individual case factors into districts by

[7] We also estimated alternative models that captured time in smaller units to see whether changes in sentencing patterns occurred more gradually. We estimated one model with dummy variables for month and year of sentence, and one model with year only. We conducted a likelihood ratio test with the month and year nested models by excluding months and years of the PROTECT Act, and then compared it with a model with our time period dummy variables. While the month and year models indicated a better fit statistically, the substantive results were nearly identical to the time period models. We were unable to conduct a nested test on the year-specific models, as we coded the time periods to the month and year of the policy changes, and many of the cases occurred midyear. The yearly models also appeared to follow a very similar pattern to the larger policy time periods. Thus, we ultimately decided to retain the larger time period models for the sake of parsimony.

[8] Although criminal history is used to calculate Guideline minimum, we were interested in additional effects above and beyond this. According to the variance inflation factor, collinearity was not an issue.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

year, including percentage of non-U.S. citizen defendants,[9] percent drug trafficking cases out of the total caseload, and percent of primary drug for trafficking cases (with crack cocaine as the reference category). We also included district-level measures of case processing, including percentage of substantial assistance, percentage of other downward departure, percentage of mandatory minimum eligible cases, and percentage of safety valve applied. We excluded these latter two measures in the mandatory minimum models.

Our state-level control variables included the drug arrest rate per 1,000 population, drug treatment admission rate per 1,000 from the Treatment Episode Data Set, the violent crime rate per 100,000 from the Uniform Crime Reports, and geographic region. Finally, we expected demographic contextual effects on sentencing practices, so controlled for percent in poverty, density of people (in hundreds per square mile), and percentage of black population. See Table 1 later for uncentered summary statistics of our key variables of interest.[10]

Before estimating the full models, we first calculated the intraclass correlation (ICC) to examine: (1) the variation in sentence length, percentage difference between the guideline minimum and actual sentence length, and use of mandatory minimums in individual cases versus between districts; and (2) the variation between mean sentence length, mean percentage difference between the guideline minimum and actual sentence length, and application of mandatory minimums *over time* within individual cases in districts versus the variation among districts. To do this, we ran an unconditional model of individual cases nested in districts, and then a second model of mean sentence length, percentage difference in sentence, and application of mandatory minimums per year in a district.

To address our research questions on variations in sentencing practices post-*Booker* and *Kimbrough/Gall*, as well as among districts with high, medium, and low proportions of drug cases, and high, medium, and low volumes of cases (measured as cases per judgeship in a district), we then divided the sample by each of these criteria and ran a series of empty models at the case and district level. This enabled us to compare variation in sentencing practices among groups based on case or district characteristics.

[9] We controlled for this at the district level because of the highly uneven distribution of immigration cases. In districts with the highest concentration of noncitizen cases, especially on the Southwestern border, drug "fast-track" programs are sometimes used, which are not available in other districts.

[10] A full table of variables is available in the online version of this article provided by Wiley.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

Table 1. Descriptive Statistics of Dependent and Key Independent Variables, fiscal years 1993–2009 (n = 280 954)

	Mean	s.d.	Min	Max
Dependent variables				
Mean sentence length	81.413	77.698	0.030	470
Percent sentence difference	85.474	64.494	0.219	16,000
Mandatory minimum eligible	0.677	0.467	0	1
Individual level predictors				
Pre-*Koon*	0.162	0.368	0	1
Koon	0.361	0.480	0	1
PROTECT Act	0.119	0.323	0	1
Booker	0.221	0.415	0	1
Kimbrough/Gall	0.137	0.344	0	1
Female	0.116	0.320	0	1
Race/ethnicity: white	0.249	0.432	0	1
Race/ethnicity: black	0.316	0.465	0	1
Race/ethnicity: Latino/Hispanic	0.412	0.492	0	1
Non-U.S. citizen	0.298	0.458	0	1
Crack cocaine	0.233	0.422	0	1
Powder cocaine	0.248	0.432	0	1
Heroin	0.076	0.264	0	1
Marijuana	0.258	0.438	0	1
Methamphetamine	0.154	0.360	0	1
Substantial assistance	0.273	0.445	0	1
Safety valve applied	0.317	0.465	0	1
District- and state-level predictors				
Percent non-U.S. citizens	32.211	22.366	0	87.076
Percent trafficking cases	39.642	11.051	8.700	75.627
Filings per judgeship rate	1.317	1.219	0.180	5.430
Median time to felony disposition	7.527	2.894	1.700	19.500
Percent substantial assistance	27.706	14.869	0	82.386
Percent mandatory minimum eligible	64.708	17.754	11.475	98.851
Percent safety valve applied	32.116	17.277	0	75.518
Region: midwest	0.175	0.38	0	1
Region: northeast	0.14	0.347	0	1
Region: south	0.499	0.5	0	1
Region: west	0.186	0.389	0	1

s.d., standard deviation.

Groups for high, medium, and low proportions of drug cases were split according to one standard deviation or greater above, 1 standard deviation within, and 1 standard deviation or lower than the mean.

Calculating the ICC for the sentence length, sentence difference, and mandatory minimum outcomes gave us the proportions of variation at the district level. The ICC is calculated as the proportion of variance at the district level over the total variance (Rabe-Hesketh & Skrondal 2005: 438): $\rho = \psi/(\psi + \theta)$ is the equation for the sentence length and percentage of sentence difference outcomes, where, ψ is the variance of the districts, and θ is the variance of the individual cases. Similarly, the conditional ICC for the dichotomous outcome of mandatory minimums in individual cases is: $\rho = \psi/(\psi + \pi^2/3)$, where $\psi + \pi^2/3$ is the total residual variance.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

We then ran a series of full random effects models, with individual cases nested in districts.[11] The general equation for the random effects model is (Rabe-Hesketh & Skrondal 2005): $Y_{ij} = \beta_1 + \zeta_{1j} + \beta_2 X_{ij} + \zeta_{2j} X_{ij} + \varepsilon_{ij}$ for i cases in j districts, where ζ_1 is a random intercept for each district, and ζ_2 is the random slope for each j district. We included random slopes for *Booker* and *Kimbrough/Gall* time periods with unstructured covariance, because we expected that these time periods could have differential impact on districts' sentencing outcomes. Because we were also interested in differences in sentencing outcomes between "high-" and "low-" mandatory minimum districts (based on the median mandatory minimum eligible rate for all districts), we conducted a Chow test to determine whether the models would be better specified as separate "high" mandatory minimum eligible districts, and "low" mandatory minimum eligible districts, and then specified separate models for each.

Results

General Trends in Drug Trafficking Case Outcomes Over Time

Actual sentencing in drug trafficking cases reflects a notable divergence from Guidelines-mandated minimum sentences across the entire time span, suggesting that the courtroom workgroups handling drug trafficking cases are collectively motivated to temper the long sentences prescribed under the Guidelines and mandatory minimum statutes. Thus, as illustrated in Figure 1, drug trafficking cases were between two and four times more likely to be sentenced below the Guideline minimum than above it, including within the Guideline range, in any given year. Fewer than 1 percent of the cases were sentenced above the Guidelines maximum, whereas about 47 percent cases were sentenced below the minimum. Overall, 85 percent of cases were sentenced at or below the Guidelines minimum, suggesting that the bottom of the range functions more as a maximum than a minimum for those actually engaged in federal adjudication and sentencing. This trend has been strong and consistent throughout the period.

Across time and case type, drug trafficking cases were sentenced to a mean total sentence that was 18 months below the Guideline minimum. The below-presumptive sentencing has actually been the

[11] In both the ICC and full HLM models, we also estimated three level full models, with cases nested in districts, nested in states, but found that results were substantially or virtually identical, so present the two-level models only. This is partially because the boundaries of more than half (26) of the states are one in the same as the districts; otherwise a given state has two to four districts within it.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

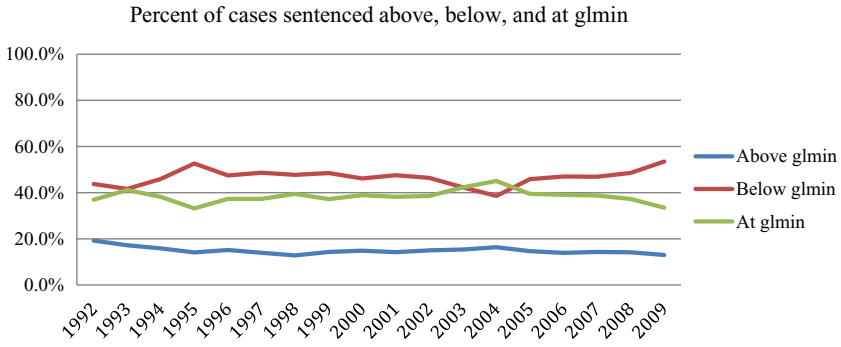

Figure 1. Drug Trafficking Sentences Relative to the Guideline Minimum.

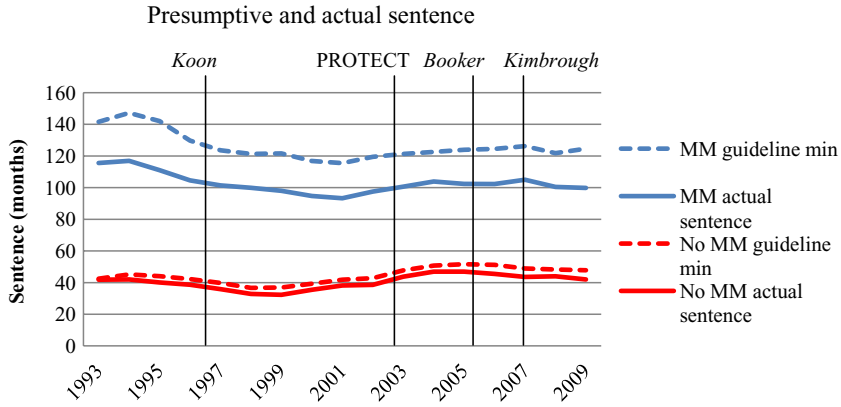

Figure 2. Actual versus. Presumptive Sentence: Mandatory Minimum vs. Guidelines-only Cases.

greatest in mandatory minimum cases. On average, final sentences in mandatory minimum cases are about 83 percent of the Guideline minimum, while nonmandatory minimum cases are sentenced at about 92 percent of the Guideline minimum (see Figure 2).

The ICCs and Sentencing Patterns Over Time and Across Districts

We first compare the variance components and the ICCs for individual cases nested in districts, and then compare the aggregated mean sentence length, mean sentence difference, and yearly proportion of mandatory minimums for each district nested in districts. As reflected in Table 2, when comparing logged sentence length for individual cases within districts versus among districts, only 11.2 percent of the variance is among districts. We then

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

Table 2. Unconditional Sentence Length (ln), Sentence Difference, and Drug Mandatory Minimum Models

	Individual Case Outcomes (n = 280,954)		Mean or Proportion Cases in a Year by District (n = 1,414)	
	Estimate	s.e.	Estimate	s.e.
Sentence length				
Intercept: β_1	4.126***	(0.034)	4.126***	(0.034)
Random effects				
District: $\sqrt{\Psi_2}$	0.320***	(0.024)	0.321***	(0.032)
Individual: $\sqrt{\theta}$	0.902***	(0.001)	0.161***	(0.010)
District ICC: ρ_2	0.112		0.799	
Sentence difference				
Fixed effects				
Intercept: β_1	86.049***	(0.896)	85.677***	−0.911
Random effects				
District: $\sqrt{\Psi_2}$	8.302***	(0.648)	8.548***	−0.696
Individual: $\sqrt{\theta}$	63.939***	(0.085)	5.917***	−0.439
District ICC: ρ_2	0.017		0.676	
Mandatory minimum				
Fixed effects				
Intercept: β_1	0.941***	(0.068)	0.704***	(0.014)
Random effects				
District: $\sqrt{\Psi_2}$	0.636***	(0.048)	0.128***	(0.010)
Individual: $\sqrt{\theta}$			0.080***	(0.005)
District ICC: ρ_2	0.089		0.719	

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. ICC, intraclass correlation; s.e., standard error.

examine the ICC for districts nested in years to observe the variation in district sentencing behaviors over time. We find that when comparing the mean logged sentence length of districts over the years, nearly 80 percent of the variance is among districts. In other words, while there is still more variation in sentence length among individual cases than districts generally, there is far greater variation in mean sentence length *between* districts than *within* districts over time.

This general pattern holds for the percentage of sentence difference and the use of mandatory minimums. For individual case outcomes, the small ICC for both percentage of sentence difference and mandatory minimums suggest that the variance occurs predominantly at the individual case level compared with the district. When examining the ICC for mean percentage of sentence difference and the proportion of mandatory minimums in districts over time, although, the proportion of variance at the level is about 68 percent, for mean percentage of sentence difference and 72 percent for proportion mandatory minimum. This suggests that there are considerable between-district differences as well as relative stability within districts over time in caseload characteristics, adjudication norms, and the metrics of sentence discounts and "going rates."

To examine the relative stability of sentencing practices across the distinct periods of sentencing policy, we ran empty models with

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

individual cases nested in districts for each time period, and compared the ICCs.[12] The proportion of variance in all sentence length outcomes among districts remains fairly stable across the periods, ranging from 11.0 to 14.2 percent for sentence length, and 12.0 to 14.9 percent for mandatory minimums. Variation still predominantly occurs at the individual case level for all outcomes in all time periods, and interdistrict variation in sentencing patterns do not increase post-*Booker*. We also ran empty models with individual cases nested in districts to examine whether caseload characteristics contributed to those patterns of stability. Results indicate that districts with relatively high numbers of trafficking cases have higher proportions of variance explained in sentence length and in use of mandatory minimums among districts compared with districts with fewer such cases. Districts with large caseloads per judge also appear more stable, in that about 21 percent of the sentence length variance is explained for high-volume districts, compared with about 10 percent for medium- and low-volume districts.

Policy Period, Case Factors, and Contextual Factors: Predicting Outcomes

In a series of full models, we specified three distinct outcome variables that might be expected to change over time in districts in response to sentencing policy reforms. Our first dependent variable is logged sentence length; the second is percentage of sentence difference; and the third is the dichotomous mandatory minimum variable. In the sentence length and sentence difference Hierarchical Linear Models (HLMs), the first model includes individual predictors only, the second model includes district and state predictors, and the third model incorporates random slopes for the *Booker* and *Kimbrough* time periods. Models 4 and 5 were estimated separately based on districts with high and low rates of mandatory minimum eligibility.[13] For all outcomes, the significant random effects for both *Booker* and *Kimbrough* indicate that both the intercept and slope varies across districts during these time periods, suggesting a diversity of sentencing practices among districts.

Regarding our findings on sentence length, prison terms were slightly longer during the PROTECT Act period than during any of

[12] Tables illustrating the remaining ICC results are in the online version of this article provided by Wiley.

[13] High-mandatory minimum districts are defined as above the mean, and low-mandatory minimum districts are below the mean. For both sets of models (sentence length and sentence difference), we conducted Chow tests by dividing our sample into "high" mandatory minimum cases and "low-" mandatory minimum districts. The joint significance test indicated that the effects are jointly different for high versus low-mandatory minimum districts, so we split the samples and estimated two separate models.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

the other policy periods, as illustrated in Table 3.[14] In comparison with the PROTECT Act, the mean sentence lengths in the *Pre-Koon, Koon*, and *Booker* periods were generally about 4–5 percent shorter, and sentences in the *Kimbrough* time period were about 8–9 percent shorter across districts. Results indicate that both the West and the Northeast sentence more leniently than the South with that effect most consistent for districts in the Northeast. Sentences in Northeast districts were 9 percent shorter than those in the South (using Model 3), all else equal.

Black and especially Latino drug defendants were sentenced more harshly than white defendants across time periods. Not surprisingly, drug type matters, where all drugs except methamphetamine predicted significantly shorter sentences compared with crack cocaine. Drug mandatory minimum eligible cases resulted in significantly longer sentences (by about 50 percent), and safety valve cases received significantly shorter sentences. Districts with more case filings per judgeship resulted in shorter sentences, suggesting that busier courts may effectively lower the "going rate" for sentences. Specifically, for an increase in 100 filings per judgeship, sentence lengths decreased about 3.5 percent.

The separate models for high and low mandatory minimums reveal substantial differences. Districts with high mandatory minimum have substantially longer sentences (about 74.23 months) than low-mandatory minimum districts (about 38.24 months). Interesting time period and regional patterns emerge as well. In the *Pre-Koon* and *Koon* eras, both high and low-mandatory minimum districts had similarly shorter sentences compared with the PROTECT Act, suggesting that the presumptive Guidelines system functioned as a brake on sentence floors in a similar manner to mandatory minimums. However, in the post-*Booker* time periods, the low-mandatory minimum districts evidenced greater decreases in predicted sentence length compared with the high-mandatory minimum districts. Specifically, sentences in high-mandatory minimum districts were 3 percent shorter during the *Booker* period, and 7 percent shorter in the *Kimbrough/Gall* period relative to PROTECT Act sentences, while sentences in low-mandatory minimum districts were 7.0 and 9.7 percent shorter in the same respective time periods.

High-mandatory minimum districts also produced more significant sentence length variance as a result of "extralegal" factors than did low mandatory districts. For instance, region mattered in the high-mandatory minimum districts only, in that sentences in

[14] The HLM tables here do not include all control variables; full tables are available in the online version of this article provided by Wiley.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

Table 3. HLM Results with Key Variables for Total Sentence Length (ln) (n = 280 954)

	Model 1		Model 2		Model 3		Model 4: High-Mandatory Min. Districts		Model 5: Low-Mandatory Min. Districts	
	Beta	s.e.	Beta	s.e.	Beta	s.e.	Beta	s.e.	Beta	s.e.
Individual level										
Pre-*Koon*	−0.037***	(0.004)	−0.046***	(0.006)	−0.043***	(0.006)	−0.040***	(0.009)	−0.040***	(0.009)
Koon	−0.042***	(0.003)	−0.036***	(0.004)	−0.029***	(0.004)	−0.028***	(0.006)	−0.030***	(0.006)
Booker	−0.048***	(0.004)	−0.045***	(0.004)	−0.052***	(0.011)	−0.031**	(0.009)	−0.069***	(0.018)
Kimbrough/Gall	−0.087***	(0.004)	−0.082***	(0.005)	−0.093***	(0.013)	−0.070***	(0.011)	−0.097***	(0.024)
Female	−0.176***	(0.003)	−0.175***	(0.003)	−0.177***	(0.003)	−0.196***	(0.004)	−0.154***	(0.005)
Black	0.022***	(0.003)	0.022***	(0.003)	0.021***	(0.003)	0.046***	(0.004)	−0.003	(0.006)
Latino/Hispanic	0.049***	(0.003)	0.049***	(0.003)	0.047***	(0.003)	0.067***	(0.005)	0.031***	(0.005)
Other race/ethnicity:	0.033***	(0.007)	0.031***	(0.007)	0.028***	(0.007)	0.021*	(0.009)	0.022	(0.011)
Non-U.S. citizen	0.042***	(0.003)	0.041***	(0.003)	0.041***	(0.003)	0.073***	(0.004)	0.024***	(0.004)
Powder cocaine	−0.013***	(0.003)	−0.014***	(0.003)	−0.014***	(0.003)	−0.031***	(0.004)	−0.008	(0.006)
Heroin	−0.014**	(0.005)	−0.019***	(0.005)	−0.019***	(0.005)	−0.042***	(0.006)	−0.002	(0.008)
Marijuana	−0.304***	(0.004)	−0.297***	(0.004)	−0.298***	(0.004)	−0.240***	(0.005)	−0.299***	(0.006)
Methamphetamine	0.093***	(0.004)	0.084***	(0.004)	0.083***	(0.004)	0.065***	(0.005)	0.092***	(0.007)
Presumptive sentence	0.006***	(0.000)	0.006***	(0.000)	0.006***	(0.000)	0.006***	(0.000)	0.007***	(0.000)
Substantial assistance	−0.588***	(0.002)	−0.59***	(0.002)	−0.59***	(0.002)	−0.611***	(0.003)	−0.57***	(0.004)
Other downward departure	−0.353***	(0.004)	−0.347***	(0.004)	−0.349***	(0.004)	−0.314***	(0.006)	−0.354***	(0.005)
Mandatory minimum eligible	0.490***	(0.003)	0.488***	(0.003)	0.489***	(0.003)	0.430***	(0.004)	0.489***	(0.004)
Safety valve applied	−0.235***	(0.003)	−0.234***	(0.003)	−0.234***	(0.003)	−0.241***	(0.004)	−0.214***	(0.004)
District- and state-level										
Percentage of non-U.S. citizens			0.000	(0.000)	−0.001*	(0.000)	0.000	(0.000)	−0.001**	(0.000)
Percentage of trafficking cases			0.000*	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Filings per judgeship rate			−0.019***	(0.003)	−0.035***	(0.003)	−0.003	(0.009)	−0.021***	(0.004)
Median time to felony disposition			−0.006***	(0.001)	−0.001	(0.001)	−0.003*	(0.001)	0.000	(0.002)
Percentage of substantial assistance			0.001***	(0.000)	0.000	(0.000)	−0.001*	(0.000)	0.001*	(0.000)
Percentage of other downward departure			0.000	(0.000)	−0.002***	(0.000)	0.000	(0.001)	−0.002***	(0.000)
Percentage of mandatory minimum eligible			0.000	(0.000)	0.000**	(0.000)				
Region: midwest	−0.029	(0.028)	0.014	(0.030)	0.016	(0.029)	−0.005	(0.028)	−0.003	(0.042)
Region: northeast	−0.19***	(0.031)	−0.136***	(0.039)	−0.098*	(0.039)	−0.167***	(0.042)	−0.084	(0.057)
Region: west	−0.155***	(0.029)	−0.081*	(0.035)	−0.050	(0.034)	−0.104**	(0.035)	−0.031	(0.047)
Intercept	3.326***	(0.017)	3.245***	(0.020)	3.237***	(0.020)	3.44***	(0.023)	3.105***	(0.029)
Random effects										
Post-Booker s.d. Ψ_{33}					0.091***	(0.008)	0.054***	(0.007)	0.115***	(0.017)
Post-*Kimbrough* s.d. Ψ_{22}					0.111***	(0.010)	0.069***	(0.008)	0.149***	(0.019)
Constant s.d. Ψ_{11}	0.099***	(0.008)	0.093***	(0.008)	0.089***	(0.007)	0.073***	(0.007)	0.111***	(0.011)
ρ32					.906***	(0.272)	.761***	(0.086)	.0953***	(0.046)
ρ31					−0.041	(0.126)	−0.080	(0.164)	−0.031	(0.121)
ρ21					−0.004	(0.126)	0.033	(0.166)	−0.062	—
Individual level s.d. θ	0.521***	(0.001)	0.520***	(0.001)	0.519***	(0.001)	0.492***	(0.001)	0.536***	(0.001)

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.
s.e., standard error; s.d., standard error.

the Northeast and West both had shorter sentences (by 16.7 and 10.4 percent, respectively) than those in the South, all else equal. Black defendants received longer sentences (by 4.6 percent) than white defendants in high-mandatory minimum districts, whereas black and white sentence lengths did not significantly differ in the low mandatory districts. Latinos in both high- and low-mandatory minimum districts received longer sentences, although that difference was more than twice as large in the high-mandatory minimum districts compared with low mandatory districts (6.7 versus 3.1 percent).

In terms of district-level predictors, the high-mandatory minimum districts are more likely to employ substantial assistance as a device for sentence reduction, whereas low mandatory districts were more likely to use other forms of departures to reduce sentences. On average, over 30 percent of cases in high-mandatory minimum districts used substantial assistance reductions, compared with about 24 percent in low-mandatory minimum districts. Other downward departures are used in 13.4 percent of cases in low-mandatory minimum districts, but only 4.7 percent in high-mandatory minimum districts. In low-mandatory minimum districts, busier districts with more filings per judgeship also predict shorter sentences.

The percentage of sentence difference models, which represent a measure of deviation from the Guidelines minimum sentence, are presented Table 4. Consistent with the sentence length models, cases were sentenced to a smaller percentage of the guideline minimum in all the time periods both before and after PROTECT Act. Sentences in the Pre-*Koon* period were about 2–5 percentage points further below the guideline minimum sentence compared with PROTECT Act sentences, 1–2 percent lower during the *Koon* period, and about 3.5–6.0 percent lower in the *Booker* and *Kimbrough* periods.

Overall, black, Latino, and "other" racial groups were sentenced 1–3 percentage points closer to the guideline minimum compared with White defendants. In substantial assistance cases, the mean sentence was 46 percentage points below the guideline minimum. The effect of other downward departures was smaller, but still substantial at over 30 percent below the minimum. Districts with higher percentages of non-U.S. citizens, and longer median times to felony disposition resulted in cases being sentenced to a smaller percentage of the guideline minimum.

As in the sentence length models, the "high-" and "low-" mandatory minimum districts differed from each other in several ways. While in both high and low-mandatory minimum districts, sentences were a smaller percentage of the guideline minimum in all periods compared with the PROTECT Act period, low-mandatory

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

Table 4. HLM Results with Key Variables only for Percentage of Sentence Difference (n = 280 954)

	Model 1		Model 2		Model 3		Model 4: High-Mandatory min. Districts[15]		Model 5: Low-Mandatory min. Districts	
	Beta	s.e.	Beta	s.e.	Beta	s.e.	Beta	s.e.	Beta	s.e.
Individual level										
Pre-*Koon*	−2.231***	(0.449)	−4.322***	(0.659)	−3.872***	(0.690)	−4.485***	(0.882)	−3.054**	(1.085)
Koon	−1.121**	(0.383)	−2.262***	(0.446)	−1.542***	(0.463)	−1.814**	(0.580)	−1.586*	(0.746)
Booker	−4.506***	(0.408)	−3.710***	(0.435)	−3.501***	(0.617)	−1.849**	(0.637)	−5.711***	(1.057)
Kimbrough/Gall	−6.287***	(0.450)	−5.616***	(0.516)	−5.703***	(0.700)	−3.480***	(0.727)	−8.851***	(1.180)
Female	−5.309***	(0.364)	−5.312***	(0.364)	−5.351***	(0.364)	−6.020***	(0.468)	−4.751***	(0.557)
Black	1.654***	(0.401)	1.556***	(0.402)	1.500***	(0.402)	1.957***	(0.475)	1.233	(0.684)
Latino	1.318***	(0.378)	1.143**	(0.379)	0.992**	(0.380)	1.790***	(0.486)	0.272	(0.589)
Other rrace/ethnicity	2.978***	(0.832)	2.779***	(0.832)	2.768***	(0.833)	0.745	(0.984)	5.633***	(1.398)
Non-U.S. citizen	0.396	(0.326)	0.472	(0.327)	0.513	(0.327)	1.165**	(0.429)	0.149	(0.493)
Powder cocaine	2.093***	(0.390)	2.075***	(0.395)	2.090***	(0.395)	1.533***	(0.447)	2.765***	(0.709)
Heroin	3.129***	(0.549)	2.947***	(0.556)	2.969***	(0.555)	1.661**	(0.639)	4.622***	(0.971)
Marijuana	0.836	(0.457)	1.061*	(0.462)	1.052*	(0.462)	1.343*	(0.586)	1.498*	(0.760)
Methamphetamine	2.214***	(0.497)	1.962***	(0.508)	1.946***	(0.508)	1.518**	(0.583)	2.696**	(0.891)
Substantial assistance	−46.114***	(0.275)	−46.309***	(0.279)	−46.324***	(0.279)	−45.900***	(0.331)	−46.818***	(0.466)
Other downward departure	−30.817***	(0.456)	−32.053***	(0.477)	−32.168***	(0.477)	−28.156***	(0.696)	−34.403***	(0.667)
Mandatory minimum eligible	−4.036***	(0.279)	−3.896***	(0.283)	−3.886***	(0.283)	−4.495***	(0.393)	−3.653***	(0.405)
Safety valve applied	−2.564***	(0.298)	−2.125***	(0.304)	−2.115***	(0.304)	−1.781***	(0.378)	−2.399***	(0.481)
District and state level										
Percentage of non-U.S. citizens			−0.077**	(0.026)	−0.055*	(0.026)	−0.061	(0.033)	−0.125**	(0.038)
Filings per judgeship rate			0.668*	(0.299)	0.197	(0.345)	0.680	(0.765)	1.307**	(0.441)
Median time to felony disposition			−0.215*	(0.101)	−0.297*	(0.120)	−0.422**	(0.134)	−0.351	(0.198)
Percentage of substantial assistance			0.039*	(0.017)	0.010	(0.018)	−0.001	(0.021)	0.031	(0.029)
Percentage of other downward departure			0.133***	(0.018)	−0.004	(0.023)	0.122*	(0.059)	−0.018	(0.028)
Percentage of mandatory minimum eligible			−0.060***	(0.017)	−0.063***	(0.018)				
Percentage of safety valve applied			−0.066***	(0.014)	−0.035*	(0.015)	−0.004	(0.025)	−0.050*	(0.020)
Region: Midwest	0.053	(1.481)	0.942	(1.667)	2.283	(1.459)	0.723	(1.618)	2.529	(1.849)
Region: Northeast	−6.646***	(1.684)	−4.627*	(2.289)	−3.764	(2.014)	−4.290	(2.421)	−1.408	(2.560)
Region: West	−6.274***	(1.576)	−3.792	(2.021)	−1.560	(1.802)	−1.378	(2.107)	−0.860	(2.119)
Intercept	106.203***	(1.146)	103.949***	(1.388)	103.765***	(1.281)	104.829***	(1.647)	103.996***	(1.823)
Random effects										
Post-*Booker* s.d. Ψ_{33}					3.636***	(0.457)	2.354	—	4.206***	(0.783)
Post-*Kimbrough* s.d. Ψ_{22}					3.900***	(0.512)	2.322	—	4.326***	(0.853)
Constant s.d. Ψ_{11}	5.164***	(0.430)	4.810***	(0.423)	4.025***	(0.402)	3.710	—	3.490***	(0.544)
ρ32					0.905***	(0.072)	0.967	—	0.859**	(0.129)
ρ31					0.399	(0.182)	0.565	—	0.343	(0.353)
ρ21					0.216	(0.189)	0.336	—	0.467	(0.317)
Individual level s.d. θ	59.870***	(0.080)	59.849***	(0.080)	59.828***	(0.080)	53.004	—	66.277***	(0.127)

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.
s.e., standard error; s.d., standard error.

[15] The standard errors for the random effects in the high-mandatory minimum districts did not converge, so we excluded them in the table.

minimum districts indicate even smaller percentages of the guideline minimum in the last two time periods. In the *Booker* time period, cases in "high-mandatory minimum" districts were 1.8 percentage points lower relative to the PROTECT Act period, compared with 5.7 percentage points lower for "low-mandatory minimum" districts. In the *Kimbrough/Gall* period, the percentage of guideline minimum was 3.5 percentage points lower for high-mandatory minimum districts, compared with nearly nine points lower in low-mandatory minimum districts.

High-mandatory minimum districts again tended to produce more variance because of "extralegal" factors than did low mandatory districts. Specifically, in high-mandatory minimum districts, Black and Latino defendants were sentenced significantly closer, by nearly 2 percentage points, to the guideline minimum sentence compared with white defendants. These disparities are in addition to the higher baseline guideline minimum sentences that Latino and especially black defendants have compared with white defendants; Latino defendants have guideline minimum sentences of 104.6 months, and black defendants have 146.6 months, compared with white defendants' mean guideline minimum sentence of 101 months. Moreover, black defendants are overrepresented in high-mandatory minimum districts, in that they represent nearly 40 percent of cases in high-mandatory minimum districts, but just 22.6 percent in low-mandatory minimum districts. Latinos, in contrast, are overrepresented in low-mandatory minimum districts. Other extralegal factors, such as being a non-U.S. citizen, also significantly impacted the amount of deviation from the guideline minimum in high-mandatory minimum districts, but not in the low ones.

Finally, we modeled mandatory minimum eligibility as a dichotomous outcome, reflected in Table 5. As we expected, mandatory minimums are significantly more prevalent in the *Koon, Booker*, and *Kimbrough/Gall* time periods compared with the PROTECT Act. By the *Kimbrough* period, mandatory minimums apply 25 percent more often than in the PROTECT Act period, suggesting that prosecutorial charging behavior may be mediating sentence outcomes, especially in the advisory Guideline periods, through the pursuit of more mandatory minimum sentences.

Substantial assistance is strongly and positively associated with mandatory minimums, and "other" departures are negatively associated, providing further evidence that different mechanisms for sentence reduction apply as a function of case type. Drug type is a consistently strong driver of mandatory minimums, where crack cocaine cases are far more likely to be eligible for mandatory minimums compared with all other drugs. Busier districts tended to have fewer mandatory minimum-eligible cases.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

Table 5. HLM Results with Key Variables for Mandatory Minimum Sentence (n = 280 954)

	Model 1			Model 2			Model 3		
	Odds Ratio	Beta	s.e.	Odds Ratio	Beta	s.e.	Odds Ratio	Beta	s.e.
Individual level									
Pre-*Koon*	0.658	−0.419***	(0.02)	0.788	−0.238***	(0.028)	0.807	−0.214***	(0.030)
Koon	1.113	0.107***	(0.017)	1.151	0.141***	(0.02)	1.146	0.136***	(0.020)
Booker	1.177	0.163***	(0.019)	1.122	0.115***	(0.02)	1.137	0.128**	(0.048)
Kimbrough/Gall	1.223	0.201***	(0.021)	1.221	0.200***	(0.024)	1.250	0.223***	(0.060)
Female	0.799	−0.224***	(0.016)	0.803	−0.220***	(0.016)	0.802	−0.221***	(0.016)
Black	0.929	−0.074***	(0.019)	0.922	−0.081***	(0.019)	0.928	−0.075***	(0.019)
Latino/Hispanic	1.203	0.185***	(0.017)	1.190	0.174***	(0.017)	1.204	0.186***	(0.017)
Other Race/ethnicity	0.911	−0.093*	(0.038)	0.907	−0.098**	(0.038)	0.927	−0.076*	(0.038)
Non-U.S. citizen	1.081	0.078***	(0.015)	1.079	0.076***	(0.015)	1.079	0.076***	(0.015)
Powder cocaine	0.807	−0.215***	(0.019)	0.803	−0.219***	(0.02)	0.800	−0.223***	(0.02)
Heroin	0.652	−0.428***	(0.026)	0.654	−0.425***	(0.026)	0.649	−0.432***	(0.026)
Marijuana	0.394	−0.932***	(0.021)	0.391	−0.939***	(0.022)	0.389	−0.945***	(0.022)
Methamphetamine	0.615	−0.486***	(0.024)	0.592	−0.525***	(0.025)	0.591	−0.526***	(0.025)
Other drug	0.028	−3.572***	(0.039)	0.028	−3.562***	(0.039)	0.028	−3.578***	(0.039)
Presumptive sentence	1.029	0.029***	(0.000)	1.029	0.029***	(0.000)	1.029	0.029***	(0.000)
Substantial assistance	1.772	0.572***	(0.014)	1.751	0.560***	(0.014)	1.758	0.564***	(0.014)
Other downward departure	0.781	−0.247***	(0.02)	0.771	−0.260***	(0.021)	0.770	−0.262***	(0.021)
District and state level									
Percentage of non-U.S. citizens				1.008	0.008***	(0.001)	1.013	0.013***	(0.002)
Percentage of trafficking cases				0.990	−0.010***	(0.001)	0.993	−0.007***	(0.001)
Percentage of powder cocaine				0.998	−0.002*	(0.001)	0.995	−0.005***	(0.001)
Percentage of heroin				0.996	−0.004**	(0.002)	0.994	−0.006**	(0.002)
Percentage of marijuana				1.000	0.000	(0.001)	0.997	−0.003**	(0.001)
Percentage of methamphetamine				1.005	0.005***	(0.001)	1.004	0.004***	(0.001)
Percentage of other drugs				0.989	−0.011***	(0.002)	0.989	−0.011***	(0.002)
Filings per judgeship rate				0.904	−0.101***	(0.013)	0.912	−0.092***	(0.016)
Median time to felony disposition				1.030	0.030***	(0.005)	1.008	0.008	(0.007)
Percentage of substantial assistance				1.004	0.004***	(0.001)	1.006	0.006***	(0.001)
Percentage of other downward departure				1.004	0.004***	(0.001)	1.006	0.006***	(0.001)
Intercept	0.985	−0.015	(0.096)	1.208	0.189	(0.12)	1.470	0.385	(0.225)
Random effects									
Post-*Booker* s.d. Ψ_{33}							1.474	0.388***	(0.036)
Post-*Kimbrough* s.d. Ψ_{22}							1.636	0.492***	(0.046)
Constant s.d. Ψ_{11}	3.310	−1.197***	(0.152)	0.318	−1.147***	(0.182)	1.714	0.539***	(0.045)
ρ32							2.230	0.802***	(0.058)
ρ31							0.888	−0.119	(0.122)
ρ21							0.843	−0.171	(0.118)

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.
s.e., standard error; s.d., standard error.

Specifically, for an increase of 100 filings per judgeship, the odds of facing a mandatory minimum decreased about 10 percent, all else equal. Finally, districts that have fewer numbers of drug cases tend to bring a relatively larger proportion of mandatory minimums in those cases.

Discussion and Conclusion

In this study, we considered how various transformations in federal sentencing law and policy impact "local" level practice. We posed three main questions, and obtained findings that speak to each. First, we asked whether there is within-district stability over time in terms of sentencing practices and outcomes. Our ICCs indicated that in any given year, individual case factors explain most of the variance observed, but districts are distinguished from each other over time, transcending the time periods. We also observed moderate regional variation in practices, with the South more punitive than the Northeast and West in some practices and outcomes. These findings suggest that while the districts that make up the federal system operate under the same formal law, the system as a whole should not be treated as a single, unified entity that responds lockstep to policy mandates.

We also asked whether organizational factors help explain district-level variance in sentencing. Our analyses suggest that they do, in that districts with proportionately larger drug trafficking caseloads, and higher caseloads per judge, demonstrate somewhat more consistency in outcomes than those with fewer drug cases and smaller criminal caseloads. These factors also created downward pressure on sentence lengths, as did the median time to case resolution. Moreover, we found that districts characterized by high rates of mandatory minimum cases differed in important ways from districts with lower rates, beyond the significant differences in sentence length. When we partitioned high- and low-mandatory minimum districts, we found that the high-mandatory minimum districts were where the racial and geographic disparities were largely occurring. In particular, black defendants were especially overrepresented in the high-mandatory minimum districts compared with the low-mandatory minimum districts, and then were over-punished relative to similar others. These findings suggest that the greater power imbalance among actors in the system, the greater risk to equality in case outcomes.

Our third set of questions asked whether the *Booker* line of cases caused sentences to vary to a greater degree from the Guidelines compared with the pre-*Booker* periods. What we found was that the short-lived PROTECT Act period stood out as the aberrational

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

sentencing period. Sentences were longer, and the gap between the Guideline minimum and actual sentences was smaller, compared with both before and after this period. These effects were especially pronounced in the "low-" mandatory minimum districts. This makes sense, given the conditions of the PROTECT Act period. It was characterized by radically restrictive and imposing top-down regulation, requiring significant oversight of sentences by appellate courts (Bailey 2004). Concurrently, the local decision-making power of U.S. attorneys was constrained at the direction of then Attorney General John Ashcroft. District-level prosecutors were required to obtain approval from the main Department of Justice for a whole range of case decisions (see Baron-Evans & Stith [2012] for details on this policy). Thus, the holistic conditions for sentencing at that period of time were distinct from the rest of the Guidelines era, and the conditions of the PROTECT Act are not likely to be (nor easily) replicated, especially in light of both the Supreme Court's recent sentencing jurisprudence, and the Attorney General's current policy that returns much decisionmaking power to district offices (Holder 2010).

We also found some support for the proposition that mechanisms used to get to normative outcomes are adapted in response to policy change. Mandatory minimum cases were more prevalent in the *Koon*, *Booker*, and *Gall* periods, compared with the PROTECT Act period. Moreover, when we compared "high" and "low" mandatory minimum groups in our sentence length and percentage of difference models, we find direct evidence of how the actual relief mechanisms play a role as a function of sentencing policy. In high-mandatory minimum districts, the percentage of substantial assistance plays a significant role in final sentence length, whereas in the low-mandatory minimum group, "other" downward departures do more sentence reduction work.

Taken together, and consistent with Ulmer et al. (2011a,b), Fischman and Schanzenbach (2012), and Rehavi and Starr (2012), our findings further call into question the Sentencing Commission's interpretation of post-*Booker* sentence outcomes. We found little evidence that judges' recently increased freedom to sentence outside of the Guidelines is the primary cause of any increases in unwarranted variations in drug trafficking sentence outcomes. Instead, we found that such problems especially emerged from how mandatory minimums were deployed. As judges have, throughout the entire period, been constrained from departing below mandatory minimums, this suggests prosecutorial behavior, including in the application of 5 K1.1 substantial assistance motions, is a critical source of such problems.

We also found evidence that when the balance of power is more evenly distributed among the "workgroup" members and when

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

trial-level courts are provided more autonomy to resolve cases, as especially demonstrated in the "low-mandatory minimum" districts post-*Booker,* sentencing norms are significantly less punitive than the law-on-the-books. This is not surprising given that drug sentences have been viewed by a wide swath of constituents, including trial-level federal judges, as too punitive under the presumptive sentencing scheme (USSC 2004). Given that the court actors who jointly adjudicate criminal matters are privy to many more facts relevant to the sentencing decision, it seems that the recalibration we observe in drug sentencing is sending an important message to policy makers that the Guidelines-prescribed sentences are on the whole too harsh and inadequate in accounting for all relevant sentencing factors.

From a policy standpoint, the existence of a dual system—an advisory Guidelines system designed to account for case and defendant factors while still facilitating cross-case uniformity, along with a mandatory minimums system that by design ignores most sentencing factors, and that is quite blunt and imprecise—provides prosecutors with an additional discretionary tool that has significant implications for sentencing outcomes. As Rehavi and Starr (2012) have demonstrated, prosecutors' pursuit of mandatory minimums at the charging stage sends cases down a distinct path for resolution, bringing with it some troubling threats to equity and justice. It is precisely this problem that contributes to racial inequality in the federal system, and that Attorney General Holder recently condemned when he directed district-level U.S. attorneys to change their charging policies so that certain low-level drug defendants, although eligible, will no longer be charged with offenses that "impose draconian mandatory minimum sentences" (Holder 2013).

Reinstating constraints on the judicial sentencing process and redistributing discretionary power back to prosecutors will neither remedy the problem of unwarranted disparities in outcomes nor prompt closer allegiance to the Guidelines. Rather than enacting additional statutory minimum sentences, as has been proposed, federal lawmakers might instead attend to the message long sent by those adjudicating cases on the ground in district courts around the nation, and adjust down the Guidelines to be in line with actual outcomes. Moreover, they might revisit statutory minimums as to their appropriateness as a sentencing tool, and even consider mechanisms for tempering prosecutorial discretion at the multiple decision-making stages in which they hold considerable or sole power—case selection, charging recommendations, and recommendations for substantial assistance departures—to better restore equilibrium in the federal criminal justice work group. It is the very nature of the prosecutor's most wielded departure weapon,

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

substantial assistance, which accounts for some of the most troubling disparities as it is governed not by the relative culpability and deservedness of the sentenced defendant, but rather on how much information she can "give" to the prosecutor in exchange for the reduction. Thus, it is not surprising that the most serious drug defendants are the ones who most benefit from this form of sentencing reduction (see USSC 2011, figures 8–11).

Our findings, more broadly, send a cautionary message about the process and prospects for legal change, especially in the context of criminal justice. Foundational to law and society is an ongoing effort to understand the various triggers, avenues, resistances, and mechanisms and roadblocks at work when formal law is implemented, which often happens in multiple and unexpected ways. In this case, we find that there has been consistent demonstrated resistance to applying the most punitive federal drug laws during the rise of the "drug war," which may be heartening to those of us who view that war as excessive and unjust. Nonetheless, the power of these laws did, in the aggregate, result in dramatically more drug cases and significantly longer sentences in the federal system (see Sutton 2013 for a similar finding at the state level). As a result, the underlying federal justice institutions are much more massive and entrenched than they were prior to the Guidelines era, making them less pliable.

Pat Carlen's (2002: 116) notion of "carceral clawback," which she defines as "the power of the prison constantly to deconstruct and successfully reconstruct the ideological conditions for its own existence," is instructive here. In the context of those institutions that generate prisoners, like the federal courts under study here, we should also expect formidable resistance to retrenchment efforts, to the extent that they are on the horizon. Local federal justice system actors, and the organizational units in which they work, are deeply invested in maintaining their legitimacy, stature, and role in the justice system, so should be expected to ideologically and operationally adapt to changing policies in order to stave off diminution. They will do so, although in ways that respond to local norms and logics, so maintenance techniques will take shape in particularized ways that diverge across districts. Ultimately, an institutional bias toward stasis should be expected to function no matter the goals of policy change, so reforms will be at least partly absorbed into entrenched practices, rendering change more symbolic than substantive (Sutton 2013). Consequently, the practical solution to punitive overindulgence—as epitomized by the federal drug sentencing practices—will likely need to cut at the core of substantive criminal law's power by fundamentally redefining culpability and recalibrating sanctions accordingly (Stuntz 2011).

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press

References

Abel, Richard L. (2010) "Law and Society: Project and Practice," 6 *Annual Rev. of Law & Social Science* 1–23.

Albonetti, Celesta (1997) "Sentencing under the Federal Sentencing Guidelines: Effects of Defendant Characteristics, Guilty Pleas, and Departures on Sentence Outcomes for Drug Offenses, 1991–1992," 31 *Law & Society Rev.* 789–822.

Bailey, Mark (2004) "Feeney's Folly: Why Appellate Courts Should Review Departures from the Federal Sentencing Guidelines with Deference," 90 *Iowa Law Rev.* 269–312.

Baron-Evans, Amy, & Kate Stith (2012) "*Booker* Rules," 160 *Univ. of Pennsylvania Law Rev.* 1631–743.

Berman, Douglas (2010) "Afternoon Keynote Address: Encouraging (and Even Requiring) Prosecutors to Be Second-Look Sentencers," 19 *Temple Political & Civil Rights Law Rev.* 249–441.

Carlen, Pat (2002) "Carceral Clawback: The Case of Women's Imprisonment in Canada," 4 *Punishment & Society* 115–21.

Cheliotis, Leonidas (2006) "How Iron Is the Iron Cage of New Penology? The Role of Human Agency in the Implementation of Criminal Justice Policy," 8 *Punishment & Society* 313–40.

Church, Thomas (1985) "Examining Local Legal Culture," 10 *Law & Social Inquiry* 449–510.

Dixon, Jo (1995) "The Organizational Context of Criminal Sentencing," 100 *American J. of Sociology* 1157–98.

Doerner, Jill K., & Stephen Demuth (2010) "The Independent and Joint Effects of Race/Ethnicity, Gender, and Age on Sentencing Outcomes in U.S. Federal Courts," 27 *Justice Q.* 1–27.

Eisenstein, James, & Herbert Jacob (1977) *Felony Justice: An Organizational Analysis of Criminal Courts*. Boston: Little, Brown.

Eisenstein, James, et al. (1988) *The Contours of Justice: Communities and Their Courts*. Boston: Little, Brown.

Engen, Rodney (2011) "Racial Disparity in the Wake of *Booker/Fanfan*: Making Sense of "Messy" Results and Other Challenges for Sentencing Research," 10 *Criminology & Public Policy* 1139–49.

Farrell, Amy, & Geoff Ward (2011) "Examining District Variation in Sentencing in the Post-*Booker* Period," 23 *Federal Sentencing Reporter* 318–25.

Feeley, Malcolm, & Sam Kamin (1996) "The Effect of 'Three Strikes and You're Out' on the Courts: Looking Back to See the Future," in Shichor, D., & D. K. Sechrest, eds., *Three Strikes and You're Out: Vengeance as Public Policy*. Thousand Oaks, CA: 135–54. Sage.

Feeley, Malcolm, & Jonathan Simon (1992) "The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications," 30 *Criminology* 449–74.

Fischman, Joshua B., & Max M. Schanzenbach (2012) "Racial Disparities under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums," 9 *J. of Empirical Legal Studies* 729–64.

Grattet, Ryken, & Valerie Jenness (2008) "Transforming Symbolic Law into Organizational Action: Hate Crime Policy and Law Enforcement Practice," 87 *Social Forces* 501–27.

Heumann, Milton (1978) *Plea Bargaining: The Experiences of Prosecutors, Judges, and Defense Attorneys*. Chicago: Univ. of Chicago Press.

Hofer, Paul J. (2007) "*United States V. Booker* as a Natural Experiment: Using Empirical Research to Inform the Federal Sentencing Policy Debate," 6 *Criminology & Public Policy* 433–60.

——— (2011) "Has *Booker* Restored Balance? A Look at Data on Plea Bargaining and Sentencing," 23 *Federal Sentencing Reporter* 326–32.

——— (2013) "The Commission Defends an Ailing Hypothesis: Does Judicial Discretion Increase Demographic Disparity?," 25 *Federal Sentencing Reporter* 311–22.

Holder, Eric (2010) *Memorandum from Eric H. Holder, Jr., Attorney General, to All Federal Prosecutors (May 19, 2010)*. Available at: http://www.justice.gov/oip/holder-memo-charging-sentencing.pdf.

——— (2013) *Remarks at the Annual Meeting of the American Bar Association's House of Delegates*. (August 12). Available at: http://www.justice.gov/iso/opa/ag/speeches/2013/ag-speech-130812.html.

Johnson, Brian, et al. (2008) "The Social Context of Guidelines Circumvention: The Case of Federal District Courts," 46 *Criminology* 737–83.

Kautt, Paula (2002) "Location, Location, Location: Interdistrict and Intercircuit Variation in Sentencing Outcomes for Federal Drug-Trafficking Offenses," 19 *Justice Q.* 633–71.

Kemshall, Hazel, & Mike Maguire (2001) "Public Protection, Partnership and Risk Penality: The Multi-Agency Risk Management of Sexual and Violent Offenders," 3 *Punishment & Society* 237–64.

Kramer, John H., & Jeffery T. Ulmer (2002) "Downward Departures for Serious Violent Offenders: Local Court "Corrections" to Pennsylvania's Sentencing Guidelines," 40 *Criminology* 897–932.

Lynch, Mona (1998) "Waste Managers? The New Penology, Crime Fighting, and Parole Agent Identity," 32 *Law & Society Rev.* 839–69.

——— (2011) "Mass Incarceration, Legal Change and Locale: Understanding and Remediating American Penal Overindulgence," 10 *Criminology and Public Policy* 671–98.

——— (2012) "Theorizing the Role of the 'War on Drugs' in US Punishment," 16 *Theoretical Criminology* 175–99.

Maddan, Sean, et al. (2012) "Sympathy for the Devil: An Exploration of Federal Judicial Discretion in the Processing of White-Collar Offenders," 37 *American J. of Criminal Justice* 4–18.

McCoy, Candace (1993) *Politics and Plea Bargaining: Victims' Rights in California*. Philadelphia: Univ. of Pennsylvania Press.

Mustard, David B. (2001) "Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts," 44 *J. of Law and Economics* 285–314.

Nagel, Ilene, & Stephen J. Schulhofer (1992) "A Tale of Three Cities: An Empirical Study of Charging and Bargaining Practices under the Federal Sentencing Guidelines," 66 *Southern California Law Rev.* 501–66.

Piehl, Anne Morrison, & Shawns Bushway (2007) "Measuring and Explaining Charge Bargaining," 23 *J. of Quantitative Criminology* 105–25.

Rabe-Hesketh, Sophia, & Anders Skrondal (2005) *Multilevel and Longitudinal Modeling Using Stata*. College Station, TX: Stata Press.

Rehavi, M. Marit, & Sonia Starr (2012) "Racial Disparity in Federal Criminal Charging and Its Sentencing Consequences," *U of Michigan Law & Econ, Empirical Legal Studies Center Paper* (12-002).

Robinson, Gwen (2002) "Exploring Risk Management in Probation Practice: Contemporary Developments in England and Wales," 4 *Punishment & Society* 5–25.

Saris, Patti (2011) *Testimony Before the, Subcommittee on Crime, Terrorism, and Homeland Security Committee on the Judiciary? United States House of Representatives*.

Savelsberg, Joachim J. (1992) "Law That Does Not Fit Society: Sentencing Guidelines as a Neoclassical Reaction to the Dilemmas of Substantivized Law," 97 *American J. of Sociology* 1346–81.

Schulhofer, Stephen J., & Ilene Nagel (1989) "Negotiated Pleas under the Federal Sentencing Guidelines: The First Fifteen Months," 27 *American Criminal Law Rev.* 231–88.

——— (1996) "Plea Negotiations under the Federal Sentencing Guidelines: Guideline Circumvention and Its Dynamics in the Post-*Mistretta* Period," 91 *Northwestern University Law Rev.* 1284–316.

Shermer, Lauren O., & Brian J. Johnson (2010) "Criminal Prosecutions: Examining Prosecutorial Discretion and Charging Decisions in U.S. Federal District Courts," 27 *Justice Q.* 394–430.

Stith, Kate, & Jose Cabranes (1998) *Fear of Judging: Sentencing Guidelines in the Federal Courts*. Chicago: Univ. of Chicago Press.

Stuntz, William J. (2011) *The Collapse of American Criminal Justice*. Cambridge, MA: Belknap Press.

Sutton, John R. (2013) "Symbol and Substance: Effects of California's Three Strikes Law on Felony Sentencing," 47 *Law & Society Rev.* 37–69.

Ulmer, Jeffery T. (2005) "The Localized Uses of Federal Sentencing Guidelines in Four U.S. District Courts: Evidence of Processual Order," 28 *Symbolic Interaction* 255–79.

——— (2012) "Recent Developments and New Directions in Sentencing Research," 29 *Justice Q.* 1–40.

Ulmer, Jeffery T., & Brian D. Johnson (2004) "Sentencing in Context: A Multilevel Analysis," 42 *Criminology* 137–77.

Ulmer, Jeffery T., & Michael Light (2011) "Beyond Disparity: Changes in Federal Sentencing after *Booker* and *Gall*?," 23 *Federal Sentencing Reporter* 333–41.

Ulmer, Jeffery T., et al. (2010) "Trial Penalties in Federal Sentencing: Extra-Guidelines Factors and District Variation," 27 *Justice Q.* 560–92.

——— (2011a) "Racial Disparity in the Wake of the *Booker/Fanfan* Decision," 10 *Criminology & Public Policy* 1077–118.

——— (2011b) "The "Liberation" of Federal Judges' Discretion in the Wake of the *Booker/ Fanfan* Decision: Is There Increased Disparity and Divergence Between Courts?," 28 *Justice Q.* 799–837.

USSC (2004) *Fifteen Years of Guidelines Sentencing: An Assessment of How Well the Federal Criminal Justice System is Achieving the Goals of Sentencing Reform.* Washington, DC: US Sentencing Commission.

——— (2006) *Final Report on the Impact of United States V. Booker on Federal Sentencing*. Washington, DC: US Sentencing Commission.

——— (2009) *An Overview of the United States Sentencing Commission*. Washington, DC: US Sentencing Commission.

——— (2010a) *2010 Federal Guidelines Manual*. Washington, DC: US Sentencing Commission.

——— (2010b) *Demographic Differences in Federal Sentencing Practices: An Update of the Booker Report's Multivariate Regression Analysis*. Washington, DC: US Sentencing Commission.

——— (2011) *Report to the Congress: Mandatory Minimum Penalties in the Federal Criminal Justice System*. Washington, DC: US Sentencing Commission.

——— (2012) *Report on the Continuing Impact of United States V. Booker on Federal Sentencing*. Washington, DC: US Sentencing Commission.

Van Slyke, Shanna, & William D. Bales (2012) "A Contemporary Study of the Decision to Incarcerate White-Collar and Street Property Offenders," 14 *Punishment & Society* 217–46.

Ward, Geoff, et al. (2009) "Does Racial Balance in Workforce Representation Yield Equal Justice? Race Relations of Sentencing in Federal Court Organizations," 43 *Law & Society Rev.* 757–806.

Wilmot, Keith, & Cassia Spohn (2004) "Prosecutorial Discretion and Real-Offense Sentencing: An Analysis of Relevant Conduct under the Federal Sentencing Guidelines," 15 *Criminal Justice Policy Rev.* 324–43.

Wu, JawJeong, & Cassia Spohn (2010) "Interdistrict Disparity in Sentencing in Three U.S. District Courts," 56 *Crime & Delinquency* 290–322.

Cases Cited

United States v. Booker, 543 U.S. 220 (2005).
Pepper v. United States, 131 S. Ct. 1229 (2011).
Kimbrough v. United States, 552 U.S. 85 (2007).
Koon v. United States, 518 U.S. 81, (1996).
Gall v. United States, 552 U.S. 38 (2007).

Mona Lynch *is Professor of Criminology, Law and Society and co-director of the Center in Law, Society and Culture at the University of California, Irvine. Her research and publications focus on the social, psychological, and cultural dynamics of criminal adjudication and punishment processes. She is the author of* Sunbelt Justice: Arizona and the Transformation of American Punishment *(Stanford Univ. Press, 2010).*

Marisa Omori *is a Ph.D. candidate in the Department of Criminology, Law & Society at the University of California, Irvine. Beginning Fall 2014, she will be an assistant professor of Sociology at the University of Miami. Her work focuses on racial inequality in the criminal justice system, and drug policy and use. She has published in* Crime and Public Policy, Theoretical Criminology, Journal of Drug Issues, *and* Crime & Delinquency.

https://doi.org/10.1111/lasr.12070 Published online by Cambridge University Press