



E-ISSN 2332-886X

Available online at

<https://scholasticahq.com/criminology-criminal-justice-law-society/>

Inter-district variation and disparities in federal sentencing outcomes: Case types, defendant characteristics, and judicial demography

Richard D. Hartley & Robert Tillyer

University of Texas at San Antonio

ABSTRACT AND ARTICLE INFORMATION

Critics of the Federal Sentencing guidelines argue that they have failed to achieve uniformity in punishment. Scholars also note that inter-district variation in sentencing hinders amelioration of unwarranted disparity; however, to date, very few studies have explored the scope and nature of inter-district disparities. The current study analyzes federal sentencing data supplemented with district level judge information obtained from the Biographical Directory of Federal Judges to explore inter-district disparities in sentencing practices and to what extent district level case types, defendant characteristics, or judicial make-up predict outcomes. Results reveal that a minimal amount of the variation in the incarceration and sentence length decisions are linked to district level factors. None of the district level factors reached significance in the incarceration model; however, significantly higher average sentences were found in districts whose defendants had higher average criminal history scores and those districts that handed out a higher percentage of sentences within the guideline range. These results reveal that while the type of case may not influence sentencing patterns at the district level, most of the inter-district variance in sentencing is due to defendant characteristics (criminal history scores) and the ways in which cases are processed (percentage of departures versus within range sentences) rather than types of cases in the district or the demography of judges on the bench. This analysis forms the foundation for a number of observations regarding sentencing outcomes using recent data from the federal judicial system.

Article History:

Received April 4, 2019

Received in revised form July 26, 2019

Accepted August 2, 2019

Keywords:

Judicial discretion, federal sentencing, inter-district disparities

© 2019 *Criminology, Criminal Justice, Law & Society* and The Western Society of Criminology
Hosting by Scholastica. All rights reserved.

Corresponding author: Richard D. Hartley, University of Texas San Antonio, 501 W. Cesar E. Chavez Blvd.,
San Antonio, TX 78207, USA
Email: richard.hartley@utsa.edu

Punishment uniformity and consistency are ideals of the administration of justice in the United States. Although legally relevant variables are the primary determinants of court outcomes, extra-legal factors such as defendant race/ethnicity and gender also influence inequalities in the application of punishment (Baumer, 2013; Spohn, 2000). Case processing factors further confound the relationship between legal and extralegal factors and sentencing decisions (Ulmer, 2012), with judicial and governmental initiated departures exerting significant influences (Hartley, Maddan, & Spohn, 2007; Johnson, Ulmer, & Kramer, 2008; Spohn & Fornango, 2009; Tillyer & Hartley, 2016).

Scholars also note that inter-district differences in sentencing practices may also undermine the goals of uniformity in punishment. Sentencing statistics confirm that the types and characteristics of cases processed, judicial characteristics, and average sentence length vary across federal districts (Johnson et al., 2008; Kautt, 2002; Waldfogel, 1991). Empirical studies suggest that *case level* legal and extra-legal factors are most impactful in understanding sentencing outcomes; however, *district level* factors also appear to be relevant (Farrell, Ward, & Rousseau, 2009; Feldmeyer & Ulmer, 2011; Johnson et al., 2008; Kautt, 2002; Ulmer, 2005; Waldfogel, 1991; Wu & Spohn, 2010). One limitation of this work is the degree of inconsistency in findings across studies. For example, there is some evidence that case characteristics (Feldmeyer & Ulmer, 2011) and case processing decisions (Ulmer, 2005) influence sentencing outcomes, while other studies found no evidence for a relationship between district level predictors and sentencing outcomes (Kautt, 2002). Additionally, there is limited work studying how district level defendant characteristics and judicial composition might impact sentencing decisions. In the case of judicial characteristics, the majority of studies examined judges at the state level leaving a gap in knowledge regarding federal judges (but see Farrell et al., 2009, for an exception). Some of the inconsistencies may also be due to the use of different methodological strategies, inclusion of varying district level predictors, and variation in number of districts and years examined. Notwithstanding the specific reasons, this area of sentencing research is still in need of development.

The current study contributes to this issue by examining the nature and extent of inter-district variation in federal sentencing decisions. Theoretically rooted in the court community and focal concerns perspectives, we explore the potential impact of *district level* case type, defendant characteristics, and judicial composition on sentence outcomes while

controlling for *case level* extra-legal, legal, and case processing characteristics. This investigation uses 2015 federal sentencing data and estimates multilevel models to identify correlates of the in/out and sentence length decisions. Initially, we summarize the existing literature on inter-district variation in sentencing outcomes and present theoretical explanations for understanding these patterns.

Literature Review

Local Courts as Community

Practically, the law is applied across a diverse federal geography by local actors resulting in differential interpretation and variability in the application of punishment, a process sometimes referred to as localized justice (Johnson et al., 2008). Thus, despite the implementation of uniform guideline-based sentencing in the federal system, concerns exist regarding whether there is “reasonable uniformity” in sentencing practices across all federal districts (Anderson & Spohn, 2010). Theoretically, attempts to explain inter-district variation often gravitate toward a focus on the organizational structure and culture of local courts (i.e., the local “contours of justice”; Eisenstein, Flemming, & Nardulli, 1988), which describes a process of flexible individual decision-making toward substantive or ideological goals external to formal legal rules (Ulmer, 1997). From this “courts as community” perspective, district level differences in sentencing practices would be shaped by formal and informal case processing norms that differ across court communities (Savelsberg, 1992; Ulmer & Johnson, 2004). In other words, even under a formal sentencing structure such as the Federal Sentencing Guidelines (FSG), inter-district variation in sentencing decisions could exist because of the local organizational and practical context (district) in which judges are embedded (Nardulli, Eisenstein, & Flemming, 1988; Ulmer, 2005). Ulmer (2012) succinctly summarizes this issue by concluding that “substantial evidence exists that what kind of sentence one gets, and the factors that predict why one gets it, in significant part depends on where one is sentenced” (p.15).

Existing empirical research generally substantiates the existence of inter-district variation. For example, Ulmer (2005) reported that sentencing outcomes were partially influenced by the negotiation of what constitutes relevant conduct to calculate the final offense level, the meaning of substantial assistance to warrant a prosecutorial departure, the effect of acceptance of responsibility on sentence reductions, and the plea agreement and/or presentence report. In other words, in the four federal districts

examined, idiosyncrasies existed in the processing and sentencing of defendants. Additional research reveals that differences in the rates of judicial and government sponsored departures contribute to inter-district sentencing disparities due to their influence on reducing base guideline sentences or circumventing mandatory minimums. For example, variability across districts in the use of departure mechanisms to mitigate sentences has been documented (Hartley, 2019; Ortiz & Spohn, 2014; Spohn, 2013), and Johnson and colleagues (2008) reported that inter-district variation in the likelihood of substantial assistance departures were due to increases in a district's caseload and racial composition. Metcalfe (2016) demonstrated that similarity and familiarity of the courtroom workgroup were both influential of case processing outcomes; specifically, sex similarity among the courtroom workgroup led to greater likelihood of a plea and shorter time to disposition whereas familiarity of the prosecutor and defense attorney more often resulted in a trial. Kim, Spohn, and Hedberg (2015), in their study of federal district courts, also reported variation in sentences due to the combination of judge and prosecutor in the case; the impact of these relationships also varied across districts.¹

The next step in understanding inter-district variation, therefore, is to identify additional factors that may be influential in producing this variation. Thus, the current study investigates whether district case type, defendant characteristics, and/or judicial composition assist in explaining inter-district variation.

The Focal Concerns Perspective

The focal concerns perspective is an alternative and complimentary approach that can be applied to understanding inter-district variation by suggesting the formal legal rules guiding the local courtroom workgroup and the substantive organizational realities of the local jurisdiction play a role in decision-making. Under this perspective, judges make assessments of offenses and offenders with regard to three focal concerns: the blameworthiness and culpability of the offender, the

need to protect the community from those who are dangerous, and the practical consequences of sentencing (Steffensmeier, Ulmer, & Kramer, 1998). Albonetti (1991) and others suggest that due to time constraints and an absence of complete information (bounded rationality) when making sentencing decisions, judges may resort to basing their decisions on stereotypes, prejudice, and past experience. Without complete information with which to make assessments of blameworthiness, protection of community, and practical constraints, judges develop a perceptual shorthand based on stereotypes and experiences that are linked to these focal concerns (Steffensmeier et al., 1998).

Other scholars have argued that bounded rationality does not characterize the federal sentencing process as there tends to be sufficient time and information available to federal judges in making sentencing determinations. However, there is still some uncertainty inherent in assessments of blameworthiness and, therefore, in predictions of future behavior of defendants (Kramer & Ulmer, 2002; Ulmer & Bradley, 2006). As such, judges are influenced by both legal (i.e., criminal history, offense seriousness) and extra-legal (i.e., defendant race/ethnicity, sex, and age) factors in deciding sentences because judges attribute group membership to statuses based on stereotypes or past experiences (Ulmer, Painter-Davis, & Tinik, 2016).

Extending the focal concerns perspective beyond the case level, the theorized judicial attributions and stereotypes may be conditionally relevant depending on the district (Johnson et al., 2008; Ulmer & Johnson, 2004). For example, Kramer and Ulmer (2009) suggested that individual courts are "distinctive organizational cultures" that develop their own unique relationship to external organizations and policies (p. 4). They argued that interpretation and adherence to sentencing guidelines, and the operationalization of focal concerns, could vary between courts under the same sentencing structure. As a result, disparities may develop in court communities that have adopted different ideas of dangerousness and risk based on stereotypes and attributions to extra-legal factors (Kramer & Ulmer, 2009). Ulmer, Painter-Davis, and Tinik (2016), for

¹ Recent studies have also identified the role of prosecutorial discretion in this process. For example, Farrell and colleagues (2009) examined the predictive utility of courtroom actor's demographic characteristics on sentencing decisions, and while only eight percent of the variance in the in/out decision could be attributed to district level measures, increases in the percentage of Black prosecutors and probation officers at the district level significantly reduced the likelihood of incarceration of Black defendants. These authors, however, found no significant findings regarding the percentage of Black judges and public defenders at the

district level. Hartley & Tillyer (2018) found that districts with higher caseloads had lower odds of cases being declined by the prosecutor or charges changed from the arrest charge. Indeed, previous research has also indicated that prosecutorial discretion may be important in understanding inter-district differences, unfortunately, measures of prosecutorial decision-making are not available in the current data, but these studies further substantiate the importance of exploring inter-district differences.

example, found significant variation in the effects of race/ethnicity on both the imprisonment and sentence length outcomes between district courts. The change in their random effects models for the Black and Hispanic male coefficients, however, revealed that this cross-district variation in the Black and Hispanic effect on sentence length is reduced when other legal and case processing factors such as offense type, criminal history, presence of departures, and pretrial detention are controlled for. The same was true for inter-district variation in imprisonment odds for black males; however, for Hispanic defendants, the between court variation in imprisonment odds could not be accounted for by the other variables in their models.

In one of the only qualitative studies to examine contextual differences in federal district courts, Harris (2018) interviewed current and former district judges and experienced criminal attorneys and found mixed results regarding whether the sentencing decisions of judges were influenced by the context of the district. The majority (79%) of interviewees agreed that the local environment had an effect on district court outcomes but disagreed on which influences were most important. Some cited judicial influences from the local legal community (i.e., active federal bar associations, culture of the courthouse), whereas others identified public opinion and local crime and incarceration rates (i.e., media coverage of crime, public outrage of certain crimes, jail and prison overcrowding) as driving judicial decisions. A consistent theme was that respondents believed that the influence of these contextual factors would vary by individual judge, such that a district level factor may affect outcomes in opposite directions depending on the judge making the decision (i.e., high crime rates in a district leading one judge to sentence more punitively and another more leniently). Based on the assertions of the courts as communities and the focal concerns perspectives, we next review the relevant empirical research examining the impact of case type, defendant characteristics, and judicial composition on court outcomes.

Inter-District Variability in Sentencing Length

Case Type

Studies examining inter-district variation reported that *case level* factors, such as the presumptive sentence, receipt of a substantial assistance departure, and pretrial custody, influenced sentence length across three federal districts in the Midwest (Wu & Spohn,

2010). Moreover, Anderson and Spohn (2010) concluded that the majority of the variation in sentence length between judges in three districts could be explained by legally relevant *case level* factors and offender characteristics (i.e., defendant's sex, employment status, and pretrial status). This theme was also echoed by Kautt (2002) who reported that legal, extra-legal, and offender characteristics differentially influenced outcomes across districts in narcotics cases.

Recent studies revealed that case type and/or case characteristics at the *district level* may also influence inter-district differences. For example, federal districts with higher proportions of drug cases and trial rates handed out longer sentences, while districts with higher per judge caseloads handed out shorter average sentences (Feldmeyer & Ulmer, 2011).² Hartley and Tillyer (2012) reported average sentences were longer in immigration cases among districts that border Mexico compared to those that do not, while Tillyer and Hartley (2016) found that average sentence length across districts was influenced by the fast-track departure rate and that districts with lower caseloads gave larger sentence discounts. Conversely, *district level predictors*, such as the substantial assistance departure rate, guideline compliance rate, and size of the population, did not significantly affect sentence severity. Sentence outcomes were also not impacted by the racial/ethnic make-up of the district or the district's unemployment rate. Collectively, the evidence regarding the influence of *district level* case type and/or characteristics on sentencing outcomes is mixed.

Defendant Characteristics

Empirical evidence also documents that that male and minority defendants (measured at the *case level*) receive harsher sentencing outcomes, net of controls for other legal and extra-legal factors and case processing factors, including departures (for more extensive reviews of the results of these studies, see Baumer, 2013; Doerner & Demuth, 2010; Johnson & Betsinger, 2009; The Sentencing Project, 2005; Spohn, 2000; Ulmer, 2012; Zatz, 2000). Drawing from the focal concerns perspective, it is also plausible that defendant characteristics at the *district level* may exert an influence on the in/out and sentence length decisions. Further, inter-district differences in how legal and extra-legal factors influence outcomes may also exist because judges may "make situational imputations about defendants' character and expected future behavior" partially due to the types of cases and

² Measures of percent Black and Hispanic in the district did not significantly affect sentence length directly, and tests for cross-level interactions (whether race and ethnicity would influence sentences

in districts with high percentages of Black and Hispanic populations) were likewise non-significant.

defendants processed in the district (Ulmer, Eisenstein, & Johnson, 2010, p. 565-566). For example, districts with high volumes of violent or drug cases may not view fraud or theft offenders as equally blameworthy and dangerous as a district with lower drug or violent crime caseloads. Likewise, some district courtroom workgroups may support the use of government sponsored and judicial departures as an efficient way to induce pleas and dispose of cases while others may not use these case processing mechanisms with frequency (Kutateladze, Adiloro, Johnson, & Spohn, 2014). In this sense, the focal concerns perspective presents a foundation for understanding the potential impact of *case type* and *defendant characteristics* on sentencing outcomes at the district level.

Judicial Composition

One additional influence on inter-district variation in federal sentencing outcomes may stem from the composition of the court itself. Some sentencing scholars argue that the nature of sentencing outcomes could be related to judicial demography, or more specifically, the diversity of judges on the bench. In other words, if more females and racial/ethnic minorities were presiding over federal cases, a different type of justice (more uniform – less disparate) would be dispensed (Goldman, 1999). Judge Bruce M. Wright epitomizes this sentiment in his statement that “most of the judges in America are white and male. The law is too pale and too male” (Washington, 1994, p. 248).

Others argue that a more diverse judiciary would not assist in ameliorating sentencing disparity citing that judges, regardless of race/ethnicity and gender, are more similar than different. In other words, the judicial selection and socialization process produces a largely homogeneous group of judges, who once on the bench, observe the rules of precedent and adhere to conventional practices (Frazier & Bock, 1982; Spohn, 2009). This is perhaps especially true of federal judges who are picked from a “very narrow stratum” of society; federal judges are likely to have gone to an Ivy League law school, were prosecutors or held lower court judgeships, or were in private practice prior to their appointment to the bench (Carp & Stidham, 1998, p. 210).

A very limited body of research has examined this issue with mixed results, and the significant majority of these studies analyzed either criminal court judges at the state level or judicial decisions in federal civil cases. For example, using Pennsylvania data, Muhlhausen (2004) found that minority judges handed out tougher sentences regardless of the defendant’s race, and Steffensmeier

and Britt (2001) found that Black male judges proscribed the harshest sentences. Spohn (1990b), using data on judges in Detroit, concluded that Black and White judges both handed out more severe punishment to Black defendants. Finally, Johnson (2014) reported that mode of conviction also matters; in Pennsylvania, minority judges handed out longer sentences in cases that went to trial but shorter sentences in plea bargained cases.

Within federal civil cases, Boyd (2016) found that African American appellate court judges were more likely to decide in favor of the plaintiff in both sex- and race-based employment discrimination cases, and Kastellec (2013), in his study of federal affirmative action cases, similarly found that Black appellate judges were more likely to vote in favor of affirmative action programs than non-Black judges. One exception to the above is a study of criminal cases in federal district court by Farrell and colleagues (2009) who found that the percentage of Black judges at the district level did not influence incarceration decisions of federal district judges.

Regarding gender, some scholars have suggested that female judges might hand out harsher sentences than their male counterparts (Collins, Manning & Carp, 2010), especially in assault and sexual assault cases and cases in which a female victim was involved (Kulik, Perry, & Pepper, 2003). Empirically, however, findings are mixed. Spohn (1990a) reported that female judges gave longer sentences in sexual assault cases in Detroit, while Johnson (2006) found no differences in outcomes between male and female judges in Pennsylvania. In a later study, Johnson (2014) found that female judges handed out shorter sentences in cases where a plea was negotiated.

Regarding federal civil cases, Boyd (2016) found that female judges were more likely than their male counterparts to vote in favor of the plaintiff in employment discrimination cases. Boyd, Epstein, and Martin (2010), however, in their study of judicial appellate court decisions in 13 different areas of civil law, found gender effects in only one type – sex discrimination cases. Female judges were more likely to decide in favor of the party alleging discrimination as opposed to their male counterparts, and where there was a female judge on a three-judge panel, the male judges on the panel were also more likely to vote in favor of the plaintiff versus an all-male panel.

Importantly, the relationship between judge and prosecutor/plaintiff is noticeably different in civil cases compared to criminal cases. The plaintiff in civil cases is generally the victim of alleged wrongdoing (e.g., discrimination) whereas in a criminal case, the prosecutor is a government official presenting a case against an alleged offender (i.e., the accused). This

distinction makes the application of findings regarding judicial composition's influence on decision-making from civil cases potentially less applicable to these relationships in criminal cases. The research on civil cases models judicial decisions related to victims alleging violations of federal law, whereas research on criminal cases models judges' decisions about appropriate punishment for those who have been found guilty of violating federal law.

Overall, the influence of judicial demography on sentencing outcomes is varied, although not extensively studied in the federal criminal justice system. Some findings, however, are consistent with the suggestion that the courtroom workgroup exerts an influence on sentencing outcomes such that disparity, in general, and inter-district variation in sentencing outcomes, specifically, may be conditioned by the characteristics of courts and the make-up of the individual courtroom actors (Ulmer, 2012). Importantly, the prevalence and causes of inter-district variation in sentencing outcomes across federal districts are still unclear, and more research is needed to explore how the characteristics of judges might condition the effects of extra-legal defendant characteristics on outcomes (Johnson, 2006; Wooldredge, 2010).

Method

Existing research on sentencing outcomes and inter-district variation collectively suggests two conclusions. First, district-level variation in sentencing outcomes exists, and while it potentially offers less explanatory power compared to case-level factors, it is nonetheless non-trivial. Second, there appears to be lack of clarity regarding variability across *all* federal districts, as many of the aforementioned studies included only a few districts in their analyses. Variability in conclusions regarding inter-district variation potentially stems from studying a limited number of districts, model misspecification, and the use of different methodologies.

The current study analyzes sentencing outcomes across 90 federal districts and includes predictors at the case and district level. The goal is to examine inter-district variation in sentencing and whether *case type*, *defendant characteristics*, or *judge characteristics* are significant predictors of outcomes. More specifically, it builds on earlier research in several ways. First, it responds to recent calls for greater exploration of inter-district processes (Ulmer, 2012), thereby contributing to an underdeveloped area of study within sentencing research. Second, it includes 90 of the 94 federal districts (we exclude districts in U.S. territories), which addresses one limitation of previous studies that examined a limited

number of districts (Anderson & Spohn, 2010; Johnson et al., 2008). Third, it assesses the overall contribution of district level processes to the overall variation in sentencing outcomes. Fourth, it tests hypotheses regarding district level processes using existing, relevant theoretical perspectives (i.e., courts as community and focal concerns). Finally, it explores three constellations of district-level factors (case type, defendant characteristics, and judge characteristics) that potentially influence sentencing outcomes.

To achieve these goals, we propose the following four research hypotheses based on previous research and the aforementioned theoretical perspectives:

H1: The odds of a prison sentence will be higher for cases processed in districts that experience higher rates of drug cases and within range cases, net of case-level predictors.

H2: The average sentence length will be longer for cases processed in districts that experience higher rates of drug cases and within range cases, net of case-level predictors.

H3: The odds of a prison sentence will be higher for cases processed in districts that handle a higher percentage of defendants who are male, minority, and possess higher criminal histories, net of case-level predictors.

H4: The average sentence length will be longer for cases processed in districts that handle a higher percentage of defendants who are male, minority, and possess higher criminal histories, net of case-level predictors.

We further propose two final hypotheses that are based on the lack of existing research in this area regarding federal judges, as well as the mixed findings of the previous research on state-level judges:

H5: The odds of a prison sentence across districts will be influenced by the judicial composition (i.e., district level gender and racial/ethnic characteristics) of district, net of case-level predictors.

H6: The average sentence length across districts will be influenced by the judicial composition (i.e., district level gender and

racial/ethnic characteristics) of the district, net of case-level predictors.

Data

Data represent federal cases processed at the district level in 2015. These data were drawn from federal cases recorded by the United States Sentencing Commission (USSC) for fiscal year 2015 (October 1, 2014 to September 30, 2015) and supplemented with information on all active district level judges in that calendar year obtained from the Federal Judicial Center, Biographical Directory of Federal Judges ($N=637$).³ These data represent cases processed in federal court where a conviction was obtained and the defendant was sentenced. Cases processed in the U.S. Territories of Guam, the Virgin and Mariana Islands, and Puerto Rico were removed, leaving data across 90 districts. Thereafter, all immigration cases were excluded because they are processed differently than other federal crimes (Ulmer et al., 2016), leaving 48,694 cases. After removal of cases due to missing information on variables of interest, 42,694 remained for analyses of the incarceration decision, and 36,956 cases were available for the sentence length models.⁴

Measures

Variables were measured at the case (i.e., L1) and district (i.e., L2) levels. The first dependent variable measures the *in/out* decision, is dichotomous, and coded 1 if the defendant received a sentence of incarceration. The second dependent variable, *sentence length*, was capped at 480 months and logged to remove the non-normality of its distribution. At the case level, several defendant characteristics were dichotomized including *male*, *White*, *Black*, *Hispanic*, *Other*⁵, and *age*. Defendant age was also squared to address the nonlinear shape of the distribution.

Defendant *criminal history* was measured using the United States Sentencing Commission's six-point scale (category one represents 0 or 1 criminal history points, and 6 represents 13 or more criminal history points).⁶ Case characteristics included the case type (i.e., *violent*, *drugs*, *white collar*, or *other*⁷), the *presumptive minimum* sentence⁸, and several dichotomized variables including a *guilty plea*, and whether the defendant was held in *custody*. Finally, dichotomous measures were created to reflect whether the sentence was *within range*, or if a departure was applied (*downward*, *upward*, *substantial assistance*, *fast track*, or *other government sponsored*).

Due to the focus on inter-district variation within the federal court system, district level variables reflect organizational factors rather than cultural or geographic factors. District level variables were therefore drawn from two sources: the aggregate L1 cases and the Federal Judicial Center database. Variables aggregated from L1 reflect the percent within district or the average within district. For example, the percent of cases resolved with an incarceration sentence (*% prison sentence*) was calculated for each district along with the *average sentence length* for those cases that resulted in an incarceration. The *number of cases*, the *number of judges*, and a measure of *caseload* were also measured. The latter was computed by dividing the number of cases by the number of judges. Case type included *% violent*, *% drug*, *% white collar*, and *% other*. *Within range* offered a measure of the percentage of cases that were sentenced within the sentencing guidelines as opposed to being processed with a departure. District level defendant characteristics measured the *% male*, *% White*, *% Black*, *% Hispanic*, and *% other race/ethnicity* that were processed through each district. The *mean*

³ Since this data set contains information on all federal judges appointed to the bench since 1789, we first selected out only District Court Judges from the 90 districts we analyzed. We then further partitioned the data utilizing the variables for service years and termination dates in order to limit the selected judges to only those that were sitting during the 2015 fiscal timeframe of the United States Sentencing Commission data utilized. These data are available at <http://www.fjc.gov/history/home.nsf/page/judges.html>. All judge information was aggregated to the district level.

⁴ The reduction of 5,738 cases from the incarceration decision to the sentence length was due to 2,696 cases that were processed with time served, 2,265 cases that did not receive a sentence, and 777 cases that possessed missing information on variables of interest.

⁵ Defendants identified as Asian, Native American, Pacific Islander, or other race/ethnicity were included in this category.

⁶ Federal probation officers calculate a defendant's criminal history score by adding points for prior convictions as part of the pre-sentence investigation report.

⁷ This category includes all cases that were not categorized as violent, drug, white collar, or immigration, and it includes crimes

such as property offenses, prostitution, pornography, gambling, and racketeering among others.

⁸ Contemporary sentencing studies include the presumptive sentence measure as a proxy for legally relevant factors (Engen & Gainey, 2000), as it controls for any mandatory minimums that trump the guideline minimum sentence (United States Sentencing Commission, 2010). The presumptive sentence was calculated in the following way: If a case possessed no mandatory minimum or did have a mandatory minimum but a safety valve provision applied, the presumptive sentence equaled the guideline minimum sentence. If the case had a mandatory minimum and no safety valve attached, and the mandatory minimum was greater than the guideline minimum, then the presumptive sentence was equal to the mandatory minimum sentence. If the case had a mandatory minimum and there was no safety valve attached, and the mandatory minimum was less than the guideline minimum, then the presumptive sentence equaled the guideline minimum sentence. This variable measures an offender's presumptive sentence under the guidelines or the minimum sentence that the judge could impose.

*defendant age*⁹ and *mean criminal history* of defendants processed in that district were also measured. Finally, judicial attributes included *mean age*, *% male*, *% White*, *% Black*, *% Hispanic*, and *% other race/ethnicity*.¹⁰

Analytic Technique

Multilevel models were estimated as the appropriate modeling technique for the nested nature of the data. Ideally, three-level models would be estimated by nesting cases within judges within districts; unfortunately, the sentencing data do not offer a judicial identifier to link the judge with a specific case. As a result, two-level models (nesting cases within districts) were estimated, and aggregate judicial demography was included as district-level predictor. Unconditional models were initially estimated to assess whether sentence length varied across districts. Level 1 predictors were subsequently entered with fixed error terms and grand mean centering¹¹ to establish a baseline for assessing the Level 2 predictors. Thereafter, L2 variables were added into the models. Given the *N* at L2 (90), the number of variables included in the models was limited. The choice of L2 variables was largely driven by previous research, theoretical justification, and avoidance of concerns of multicollinearity. All results reported use robust standard errors and tests for multicollinearity demonstrated no values above 3.0.

Results

Descriptive statistics indicate that 88% of all cases resulted in a prison sentence (i.e., in/out decision; see Table 1).¹² Of the defendants who received between 0.3 and 480 months' imprisonment, defendants received roughly 67 months of imprisonment on average (see Table 1). Defendants were mostly male with a slightly higher percentage of males in the sentence length group (86%) compared to the in/out decision (83%). White, Black, and Hispanic

defendants comprised between 29% and 34% of the cases, with the remaining cases reflecting the other race/ethnicity group. Defendants were roughly 37 years of age in both samples, and defendants sentenced to a prison term possessed a slightly higher criminal history (2.6% versus 2.4%).

Drug cases represented the plurality with a slightly higher representation in the sentence length sample (49%) compared to the in/out decision (45%). Violent cases represented roughly 5% in both groups. Cases ranged noticeably on the presumptive minimum, with a higher average in the sentence length cases (86%) compared to the in/out decision (76%). The large majority of cases involved a guilty plea (95-96%). Defendants held in custody were slightly more pronounced in the sentence length sample (74%) compared to the in/out decision (65%). Roughly 40% of the cases were sentenced within range in both samples, with downward (25-26%) and substantial assistance (18%) departures the most common reason for adjustments.

At the district level (see Table 2), the percentage of cases resulting in a prison sentence ranged from a low of 55% to a high of 96% across the districts with an average of 85%. Sentence length ranged from a low of 18.72 months to a high of 107.86 with an average of 67.63 months across the 90 districts. The number of cases and number of judges also demonstrated wide variability with roughly 541 cases, on average, processed in a district, an average of 7.5 judges per district, and a caseload that ranged from a low of 13.78 to a high of 283.69 with an average of 79 cases per judge. Drug cases dominated the caseload in most districts. On average, 40% of district caseload involved drugs, but this varied from nearly 80% of the caseload in some jurisdictions to as little as 16% in others. White collar cases were the next most common (20%), followed by other offenses (17%), with violent cases the least frequent (5%). Districts, on average, meted out within range sentences in 41% of the cases, but this ranged from a low of 12% of cases to a high of 70% of cases.

⁹ The defendant age squared was used to calculate the district level mean defendant age.

¹⁰ Judges identified as Asian, Native American, Pacific Islander, or other race/ethnicity were included in this category.

¹¹ Models were also estimated with group mean centering of variables that were aggregated from Level 1. No substantive differences were found in the results.

¹² Note that 6.3% of these cases were processed as time served.

Table 1: Case Descriptives

			In/Out Decision (N = 42,694)		Sentence Length (N = 36,956)	
	Min	Max	Mean	S.D.	Mean	S.D.
In/Out Decision	0.00	1.00	0.88	0.32	--	--
Sentence Length	0.03	480.00	--	--	66.72	68.40
Logged Sentence Length	-3.51	6.17	--	--	3.62	1.40
Defendant Characteristics						
Male	0.00	1.00	0.83	0.37	0.86	0.34
White	0.00	1.00	0.34	0.47	0.32	0.47
Black	0.00	1.00	0.29	0.46	0.30	0.46
Hispanic	0.00	1.00	0.31	0.46	0.33	0.47
Other	0.00	1.00	0.05	0.23	0.05	0.22
Age (In/Out) (Sentence Length)	16.00 16.00	87.00 86.00	37.42	11.74	36.76	11.24
Age Squared (In/Out) (Sentence Length)	256.00 256.00	7569.00 7396.00	1538.30	961.75	1477.87	936.46
Criminal History	1.00	6.00	2.44	1.77	2.60	1.82
Case Characteristics						
Violent	0.00	1.00	0.05	0.25	0.05	0.22
Drugs	0.00	1.00	0.45	0.48	0.49	0.50
White Collar	0.00	1.00	0.20	0.45	0.17	0.38
Other	0.00	1.00	0.14	0.34	0.12	0.33
Presumptive Minimum	0.00	480.00	75.55	81.00	86.45	82.19
Guilty Plea	0.00	1.00	0.96	0.20	0.95	0.21
Custody	0.00	1.00	0.65	0.48	0.74	0.44
Within Range	0.00	1.00	0.40	0.49	0.41	0.49
Downward	0.00	1.00	0.26	0.44	0.25	0.43
Upward	0.00	1.00	0.02	0.15	0.03	0.16
Substantial Assistance	0.00	1.00	0.18	0.39	0.18	0.38
Fasttrack	0.00	1.00	0.03	0.16	0.03	0.18
Other Government	0.00	1.00	0.10	0.30	0.10	0.30

Districts largely processed cases involving defendants who were male (83%), non-White (60%), and in their late 30s on average (37.66). There was considerable variability in defendant race/ethnicity, as Whites comprised 73% of defendants in one district and 7% in another district. Defendant criminal history also varied between districts with a low of 1.78 to a high of 3.19. Judge characteristics did not vary noticeably across districts. For example, judges, on average, were 61 years of age with a small standard deviation, and districts were predominately comprised of judges who were male (70%) and White (82%). Of note, minority judges were fairly well represented in some districts,

as the range for Hispanic judges was 0 to 57% while for Black judges it was 0 to 44%.

An unconditional model was initially estimated to identify variation in the in/out decision and sentence length across districts. Results from the intra-class correlation (ICC) indicate that 94.0% of the variance in the in/out decision and 95.7% of the variation in sentence length rests at the case level (see Tables 3 & 4). Thus, the considerable majority of variance in the in/out decision and sentence length rests within the case; nonetheless, there was a statistically significant amount of inter-district difference to explore in each model.

Table 2: District & Judge Descriptives (N= 90)

	Min	Max	Mean	S.D.
% Prison Sentence	0.55	0.96	0.85	0.07
Average Sentence Length	18.72	107.86	67.63	14.61
# of Cases	81.00	3,688.00	541.04	539.54
# of Judges	2.00	29.00	7.46	5.79
Caseload	13.78	283.69	78.88	43.91
Case Characteristics				
% Violent	0.01	0.34	0.05	0.42
% Drugs	0.16	0.78	0.40	0.13
% White Collar	0.04	0.40	0.20	0.07
% Other Offense	0.04	0.48	0.17	0.06
% Within Range	0.12	0.70	0.41	0.13
Defendant Characteristics				
% Male	0.72	0.91	0.83	0.05
% White	0.07	0.73	0.40	0.15
% Black	0.02	0.71	0.32	0.18
% Hispanic	0.01	0.83	0.21	0.18
% Other Race/Ethnicity	0.00	0.61	0.07	0.11
Average Defendant Age	31.23	41.87	37.66	1.92
Average Criminal History	1.78	3.19	2.53	0.35
Judge Characteristics				
Average Judge Age	51.80	72.50	60.50	3.38
% Judge Male	0.33	1.00	0.70	0.16
% Judge White	0.25	1.00	0.82	0.19
% Judge Black	0.00	0.44	0.11	0.13
% Judge Hispanic	0.00	0.57	0.05	0.11
% Judge Other Race/Ethnicity	0.00	0.75	0.02	0.09

Table 3: Unconditional model of In/Out Decision (L1 = 42,694; L2 = 90)

Random Effect	Variance	SD	X ²	Percentage of Total Variance
Prison Sentence (Level 2)	0.213	0.461	934.82****	6.0 ¹³

NOTE: *p≤0.10; **p≤0.05; ***p≤0.01; ****p≤0.001

Table 4: Unconditional model of Sentence Length (L1 = 36,956; L2 = 90)

Random Effect	Variance	SD	X ²	Percentage of Total Variance
Mean Sentence Length (Level 2)	0.084	0.289	1,977.195****	4.3
Within Case (Level 1)	1.866	1.366		95.7

Table 5, Model 1 summarizes a multilevel model examining the in/out decision with L1 predictors only. Extra-legal variables were highly predictive of a prison sentence with male defendants experiencing 1.5 times higher odds of receiving this

outcome, while Black and Hispanic defendants also possessed 1.3 and 1.6 higher odds of receiving a prison sentence, respectively, compared to White defendants. Defendants with a higher criminal history score also were more likely to be sentenced to prison. Legal (or

¹³ The interclass correlation coefficient calculated using the formula suggested by Snijders and Bosker (1999) ($\tau_{00} / (\tau_{00} + \pi^2/3)$) that is appropriate for dichotomous dependent variables and indicated that

6.0% of the variation in the likelihood of a prison sentence rests at the district level.

case) characteristics also increased the odds of receiving a prison sentence. For example, cases processed based on a violent offense possessed higher odds of a prison term compared to drug cases. Elevated odds of a prison sentence were also present in cases with higher presumptive minimums and most impactful for cases in which the defendant was held in custody prior to trial. The odds of a prison sentence were lower in cases involving non-drug, non-violent, and non-white collar offenses and in cases that involved a guilty plea.

The full model including district level predictors is presented in Table 5, Model 2. These variables were chosen to explore whether the case

type, defendant characteristics, and judicial composition of the districts was related to the percentage of a prison sentence issued in a district. Note, to limit the number of variables included at the district level due to the relatively small n at L2 ($n=90$), only eight district level variables were included in the model. In addition, multicollinearity concerns resulted in discarding the average defendant age from the model. Estimation of this model revealed that the case level variables were virtually unchanged from Model 1. Importantly, no district level predictors included in the model were statistically related to the percentage of a prison sentence handed out in each district.

Table 5: Multilevel In/Out Models (L1 = 42,694; L2 = 90)

Fixed Effect	Model 1			Model 2		
	Coeff.	S.E.	Odds Ratio	Coeff.	S.E.	Odds Ratio
Intercept	5.79***	0.31	--	5.78***	0.31	--
Defendant Characteristics						
Male	0.43***	0.05	1.54	0.43***	0.05	1.54
Black	0.25***	0.06	1.28	0.25***	0.06	1.28
Hispanic	0.44***	0.09	1.55	0.43***	0.09	1.54
Other	0.17	0.10	--	0.17	0.10	--
Age Squared	-0.00***	0.00	0.99	-0.00***	0.00	0.99
Criminal History	0.25***	0.03	1.29	0.25***	0.03	1.29
Case Characteristics						
Violent	0.37*	0.16	1.45	0.38*	0.17	1.46
White Collar	0.07	0.08	--	0.07	0.07	--
Other Offenses	-0.34***	0.09	0.71	-0.34***	0.09	0.71
Presumptive Minimum	0.05***	0.00	1.06	0.05***	0.00	1.06
Guilty Plea	-0.67***	0.12	0.51	-0.67***	0.12	0.51
Custody	2.89***	0.13	18.10	2.89***	0.13	17.95
District Characteristics						
Case Characteristics						
Caseload	--	--		0.00	0.00	--
% Drug Cases	--	--		0.51	0.46	--
Defendant Characteristics						
% Male	--	--		-0.66	1.29	--
% White	--	--		0.42	0.53	--
Average Criminal History	--	--		-0.10	0.17	--
Judicial Composition						
Average Judge Age	--	--		-0.00	0.02	--
% Judge Male	--	--		0.32	0.27	--
% Judge White	--	--		-0.38	0.39	--
Random Effect	Var.	SD	X²	Var.	SD	X²
Mean Sentence Length	0.196	0.443	508.69***	0.204	0.452	542.31***

NOTE: * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$; All models used Robust Standard Errors (RSE); All L1 variables were fixed due to our interest in L2 effects.

Reference Groups: White Defendant, Drug Case

Table 6 summarizes a case level model only of sentence length (Model 1) and a full model that includes district level predictors (Model 2). Statistically significant predictors at the case level include male defendants and those with larger criminal history scores who received longer sentences. There were no statistical differences in sentence length for most race/ethnic groups with the exception of Asian, Native American, or Pacific Islander (other) defendants who received statistically significant, but negligibly shorter sentences compared to White

defendants. The finding that Black and Hispanic defendants received sentences comparable to their White counterparts is a novel finding and one that is contradictory to most previous research on sentencing – this will be discussed in more detail in the Discussion. Several legal (or case) variables influenced sentence length. For example, cases prosecuted for non-violent and non-drug offenses received shorter sentences (compared to violent and drug cases), and cases involving downward, substantial assistance, fast track, or other government

sponsored departures experienced shorter sentences compared to cases sentenced within range. Conversely, cases involving higher presumptive minimums, a defendant held in custody, and upward departures all resulted in longer sentences.

Model 2, Table 6 reports the simultaneous effect of L1 and L2 variables on sentence length. The effects of the case level variables did not change with the addition of the L2 district predictors; however, a

couple of the district level variables exerted a statistically significant influence on sentence length. Districts that processed a higher percentage of cases within range produced longer average sentences, and districts with higher average defendant criminal history scores also resulted in longer average sentences. No judicial characteristics were related to average sentence length across the 90 districts.¹⁴

Table 6: Multilevel Sentence Length Models (L1 = 36,956; L2 = 90)

Fixed Effect	Model 1		Model 2	
	Coeff.	S.E.	Coeff.	S.E.
Intercept	3.63***	0.02	3.63***	0.02
Defendant Characteristics				
Male	0.21***	0.03	0.21***	0.03
Black	-0.01	0.02	-0.01	0.02
Hispanic	-0.00	0.02	-0.00	0.02
Other	-0.09*	0.04	-0.09*	0.04
Age Squared	0.00*	0.00	0.00*	0.00
Criminal History	0.04***	0.01	0.04***	0.01
Case Characteristics				
Violent	0.01	0.04	0.01	0.04
White Collar	-0.24***	0.04	-0.24***	0.04
Other Offenses	-0.26***	0.03	-0.26***	0.03
Presumptive Minimum	0.01***	0.00	0.01***	0.00
Guilty Plea	0.00	0.04	0.00	0.04
Custody	0.54***	0.06	0.54***	0.06
Downward	-0.32***	0.03	-0.31***	0.03
Upward	0.54***	0.03	0.54***	0.03
Substantial Assistance	-0.62***	0.07	-0.62***	0.06
Fast Track	-0.61**	0.18	-0.60**	0.18
Government Sponsored	-0.39***	0.03	-0.38***	0.03
District Characteristics				
Case Characteristics				
Caseload	--	--	0.00	0.00
% Drug Cases	--	--	-0.17	0.18
% Within Range	--	--	0.54***	0.14
Defendant Characteristics				
% Male	--	--	-0.56	0.53
% White	--	--	-0.11	0.15
Average Criminal History	--	--	0.15*	0.06
Judicial Composition				
Average Judge Age	--	--	0.01	0.01
% Judge Male	--	--	0.18	0.12
% Judge White	--	--	-0.09	0.13
Random Effect	Var.	SD	Var.	SD
Mean Sentence Length	0.038	0.195	0.030	0.174
		1,700.54***		1,183.66***

NOTE: *p≤0.05; **p≤0.01; ***p≤0.001; All models used Robust Standard Errors (RSE); All L1 variables were fixed due to our interest in L2 effects.

Reference Groups: White Defendant, Drug Case, Within Range

¹⁴ Note: this model did not include average defendant age due to multicollinearity concerns.

Discussion

The federal sentencing guidelines were promulgated to increase uniformity and consistency in punishment across the federal court system. Since their implementation, however, critics argue that districts develop informal rules and procedures for processing cases based on local differences that result in inter-district disparities for sentencing outcomes. Examination of federal sentencing data reveals a limited amount of variability across districts in the percentage of cases resolved with a prison sentence, while the variation in average sentence length across districts is more pronounced with the average sentence length across districts ranging from a low of roughly 18 months to a high of over 107 months. Moreover, the multilevel models reveal specific predictors of inter-district variation in sentence length, but not the in/out decision. This analysis forms the foundation for a number of observations regarding sentencing outcomes using recent data from the federal judicial system.

While the focus of this study was on inter-district variation, the results revealed valuable information regarding case level predictors' relationship with sentencing outcomes. For example, the majority of variation in the decision to assign a prison sentence (94.0%) and the length of that sentence (95.7%) predominately rests at the case level and is informed by defendant and case characteristics. The in/out model largely produced findings consistent with long-understood expectations for the influence of extra-legal and legal factors, whereas the sentence length model produced unexpected findings regarding the experience of minority defendants. While the "usual" predictors were again found to be influential (e.g., defendant gender, criminal history score, case type, and departures among others), no statistical difference was found between Black, Hispanic, and White defendants in sentence length. Admittedly, this is difficult to reconcile given the preponderance of previous research outlining the disparity in sentence length among minority defendants (e.g., Spohn, 2000). However, reviews of the body of race and sentencing research reveal that racial disparities disadvantaging Blacks and Hispanics were greater for the incarceration decision than the sentence length determinations (Baumer, 2013; Spohn, 2000). Baumer (2013) also reports that whether race effects are found in sentencing studies is dependent on the types of samples and crimes analyzed as well as the analytic models utilized.

In that vein, we offer a few possible explanations for this result. First, these USSC data are more recent than many current studies and may indicate that judges have become more attentive to

bias in decision-making, at least with respect to sentence length decisions, and are meting out more comparable punishments for minority defendants due to their awareness of statistics and social commentary regarding harsh treatment of defendants (particularly Black defendants) in the federal system. This awareness may be coupled with recent Supreme Court decisions (see *Gall v. United States*, 2007; *Kimbrough v. United States*, 2007; *United States v. Booker*, 2005) rendering the federal sentencing guidelines advisory to allow judges freedom to use their discretion to target long-standing disparities in the system. Finally, the trend in the federal criminal justice system away from the war on drugs and toward a war on immigration, as well as differences in federal prosecutorial priorities for certain types of drugs (i.e., crack-cocaine), and the conflation of the mandatory minimum trigger amount associated with drugs that minorities, especially Blacks, were more likely to be prosecuted for might also be a contributing factor to the non-significant findings regarding race and ethnicity in our sentence length model. Obviously, this is speculative, but if this pattern holds in future years to become a trend, it may signal changes in the processing of cases in the federal system and subsequent influence on prosecutorial and judicial decisions with respect to court outcomes and extra-legal disparity regarding defendant race and ethnicity.

Beyond the case level implications, the results present limited support for the district level influence of case type, defendant characteristics, and judicial composition on sentencing outcomes. For example, district level case type (drug case or within range case) was not significantly related to the in/out decision across districts, and, therefore, our findings provide no support for hypothesis 1. This non-significant result was also found for the district level percent of drug cases in our sentence length model, and, therefore, we find only partial support for hypothesis 2. This result was somewhat surprising as variation exists across districts in the types of offenses they processed. For example, drug cases comprised 78% of all cases in one district, while only 16% of all cases in another district. Moreover, this is inconsistent with previous research documenting that offense type affects sentence length under the federal sentencing guidelines (Feldmeyer & Ulmer, 2011; Ulmer et al., 2010). Future research might seek to uncover the effects of district case type on outcomes by partitioning districts into circuits or contiguous states. Districts in the southwestern United States, for example, process the majority of drug and immigration cases, and previous research has shown differences in average sentence lengths between districts that border Mexico and the other districts

(Hartley & Tillyer, 2012) and between districts in adjacent states (Anderson & Spohn, 2010).

Of note, however, districts that processed more cases within sentence range experienced higher average sentence lengths indicating that while the type of case may not influence sentencing patterns at the district level, other case characteristics are influential. This finding is consistent with predictions offered by the court as a community perspective (Nardulli et al., 1988; Savelsberg, 1992; Ulmer, 2005; Ulmer & Johnson, 2004), which predicts that the federal sentencing guidelines may be practically applied in various ways across the federal districts. In districts where a larger percentage of cases are sentenced within the guideline range, lengthier average sentences are produced. While we have limited information on case characteristics beyond case type and the use of departures to fully understand how this constellation of factors may influence sentence length across districts, additional study of these processes may reveal that the types and characteristics of cases processed through a district do exert an influence on the aggregate patterns of judicial decision-making. Additionally, judicial decisions to sentence within the guideline range might also be influenced by other contextual and organizational factors at the district level such as local legal cultures, public opinion, and crime and incarceration rates, as well as the case processing norms of the local courtroom workgroup (Harris, 2018; Metcalfe, 2016). To the extent that some of these contextual differences are driving the percentage of within-range sentences at the district level, we cannot further disaggregate the meaning of this finding. Future research should continue to include other district-level contextual factors in models of sentencing outcomes.

Similar to case type, the district level defendant characteristics demonstrated no impact on the percentage of cases resulting in a prison sentence across districts and only limited influence on the average sentence length across districts. In particular, only districts processing defendants with higher average criminal history scores produced longer average sentences. We, therefore, also find no support for hypothesis 3 and only partial support for hypothesis 4. While uniformity in sentence length across districts is the ideal, part of the explanation for an inter-district disparity being due to a legal factor such as the average criminal history score of defendants is not as concerning compared to the discovery of extra-legal defendant factors such as percent male or minority. A caveat to this is that average criminal history scores at the district level could also be dependent on the types of cases prosecuted in the district (e.g., the percentage of violent and drug cases versus immigration cases) or

local organizational factors related to case processing norms of the courtroom workgroup (e.g., how final criminal history points are calculated, and/or the defense attorney's willingness to challenge that calculation and which final determination a judge accepts). Recent research has also revealed that extra-legal factors at the *case level* interact with legally relevant factors to influence court outcomes (Rodriguez, Curry, & Lee, 2006; Tillyer, Hartley, & Ward, 2015). Therefore, future research should explore whether these types of interactions also operate at the *district level* to further explore the nuanced ways in which disparities might continue in criminal court processing.

Finally, no judicial characteristics were influential on the in/out decision or sentence length at the district level. These findings, therefore, do not provide support for hypotheses 5 and 6. This finding also runs contrary to some previous research at the state level (Muhlhausen, 2004; Spohn, 1990a; Steffensmeier & Britt, 2001) but is consistent with other evidence (Johnson, 2014) and commentary suggesting that judicial characteristics of federal judges have a limited (Johnson, 2006) and/or mixed effect (Wooldredge, 2010) on sentencing outcomes. The inconsistency of earlier literature with regard to judicial effects may be rooted in differences between the federal and state systems. While judges as a whole are rather homogeneous compared to the broader population, it is conceivable that state judges are more heterogeneous and exert greater discretion compared to federal judges resulting in more influence on sentencing outcomes. While no consistent effect of judicial demography was discovered, it may be useful for future research to measure the judicial philosophy of federal judges related to the purposes of punishment and/or thoughts about whether the guidelines are achieving those purposes through qualitative interviews to allow a more developed understanding of their decision-making processes (Anderson & Spohn, 2010). Overall, our results reveal limited explanatory power for district level case type, defendant characteristics, or judicial composition in understanding sentencing outcomes. This raises questions about what other factors influence decision-making practices at the district level. Previous research included other measures of jurisdictional level factors, including unemployment rates, violent crime rates, percent minority population, and political context, with minimal or no effects (Farrell et al., 2009; Fearn, 2005; Feldmeyer & Ulmer, 2011; Johnson et al., 2008; King, Johnson, & McGeever, 2010; Ulmer & Johnson, 2004). Given these results, we argue that researchers should endeavor to measure other characteristics and processes of the federal courtroom workgroup at the district level in order to

explore possible influences of the unexplained variance. Measures that attempt to capture the empirical reality of organizational and caseload pressures as well as the various ideals, goals, and attitudes and experiences of, and among, the court actors (AUSAs, Public Defenders and Judges) are difficult to attain, but critical.

Recent research focusing on prosecutorial discretion at the case level may offer additional value if applied to the district level and assess how prosecutors as a group differ across federal districts and how this might influence case outcomes in these jurisdictions. Recent research (Hartley & Tillyer, 2018; Johnson, 2018; Lynch, 2018) has uncovered inter-district variation in prosecutorial decisions about the selection of cases for prosecution, charge reduction, and declination. If the substantive rationalities of local court communities affect interpretation and application of formal legal rules, inclusion of these variables would be necessary to more fully test the court community perspective. Notwithstanding the aforementioned findings, limitations exist. For example, judges were not able to be nested within cases (e.g., Wooldredge, 2010). Ideally, a tri-level model consisting of cases nested within judges nested within districts would offer the most comprehensive approach to understanding sentencing outcomes. While we lack the ability to draw conclusions about specific judges, we were able to assess district level processes that conform to our stated interest of studying inter-district variation in sentence length. Also of note, the federal sentencing data present only limited measures of case processing, defendant characteristics, and overall decision-making. Ultimately, qualitative data from specific criminal justice actors would enhance our understanding of these processes and increase our ability to test theoretical explanations. We attempted to offset this limitation by introducing data on judicial characteristics, but admittedly, these efforts still leave gaps in our ability to estimate comprehensive models.

Conclusion

In conclusion, districts account for a limited, but not insignificant, portion of the overall variance in sentence outcomes. This variation is partially explained by case characteristics (i.e., percent of cases processed within guideline range) and the type of defendant processed through the district (e.g., average defendant criminal history score). Given the small variation in sentencing outcomes attributable to the district, it appears that judges are not that far apart in their determinations of appropriate sentences under the Federal Sentencing Guidelines. When differences arise, it appears they are at least partially explained by

legal factors that presumably we would expect to have an influence over sentencing processes. Potentially because they have been working with sentencing guidelines now for over a quarter century, it may be that judges are collectively cognizant of what punishments are appropriate for specific offenses, or at least the floors and ceilings of what is constitutionally acceptable, and thus district level effects (case type, defendant type, and judge composition) are relatively small.

Related, the court community perspective may offer guidance in understanding how aggregate processes operate at the district level to produce effects on sentencing outcomes. Although more specific measures of district level court and individual judge contexts are necessary to expand this line of research, our analysis provides some broad ideas about potentially important jurisdictional characteristics and district specific contexts that influence final sentencing outcomes. Models that account for additional factors measuring case-specific and organizational contexts of court decision-making will assist in advancing our knowledge of the nature of federal punishment, and the important legal, extra-legal, and contextual determinants of judicial sentencing practice

References

- Albonetti, C. A. (1991). An integration of theories to explain judicial discretion. *Social Problems*, 38, 247–266. doi: [10.2307/800532](https://doi.org/10.2307/800532)
- Anderson, A. L., & Spohn, C. (2010). Lawlessness in the federal sentencing process: A test for uniformity and consistency in sentencing outcomes. *Justice Quarterly*, 27, 362–393. doi: 10.1080/07418820902936683
- Baumer, E. (2013). Reassessing and redirecting research on race and sentencing. *Justice Quarterly*, 30, 231–261. doi:10.1080/07418825.2012.682602
- Boyd, C. L. (2016). Representation on the courts? The effects of trial judges' sex and race. *Political Research Quarterly*, 69, 788–799. doi: 10.1177/1065912916663653
- Boyd, C. L., Epstein, L., & Martin, A. D. (2010). Untangling the causal effects of sex on judging. *American Journal of Political Science*, 54, 389–411. doi: 10.1111/j.1540-5907.2010.00437.x
- Carp, R. A., & Stidham, R. (1998). *Judicial process in America*. Washington, DC: Congressional Quarterly.

- Collins, P., Mannign, K., & Carp, R. (2010). Gender, critical mass, and judicial decision making. *Law & Policy*, 32, 260-281.
- Doerner, J. K., & Demuth, S. S. (2010). The independent and joint effects of race/ethnicity, gender, and age on sentencing outcomes in US federal courts. *Justice Quarterly*, 27, 1-27. doi: 10.1080/07418820902926197
- Eisenstein, J., Flemming, R. B., & Nardulli, P. F. (1988). *The contours of justice: Communities and their courts*. Boston, MA: Little, Brown.
- Engen, R.L. & Gainey, R.R. (2000). Modeling the effects of legally relevant and extralegal factors under sentencing guidelines: The rules have changed. *Criminology*, 38, 1207-1230.
- Farrell, A., Ward, G., & Rousseau, D. (2009). Race effects of representation among federal court workers: Does black workforce representation reduce sentencing disparities? *The Annals of the American Academy of Political and Social Science*, 623, 121-133. doi: [10.1177/0002716208331128](https://doi.org/10.1177/0002716208331128)
- Fearn, N. E. (2005). A multilevel analysis of community effects on criminal sentencing. *Justice Quarterly*, 22, 452-487. doi: [10.1080/07418820500364668](https://doi.org/10.1080/07418820500364668)
- Feldmeyer, B., & Ulmer, J. T. (2011). Racial/Ethnic threat and federal sentencing. *Journal of Research in Crime & Delinquency*, 48, 238-270. doi: [10.1177/0022427810391538](https://doi.org/10.1177/0022427810391538)
- Frazier, C. E., & Bock, E. W. (1982). Effects of court officials on sentence severity: Do judges make a difference? *Criminology*, 20, 257-272. doi: 10.1111/j.1745-9125.1982.tb00460
- Gall v. United States, 552 U.S. 38 (2007)
- Goldman, S. (1979). Should there be affirmative action for the judiciary? *Judicature*, 62, 488-494.
- Harris, S. (2018). In their words: The sentencing decisions of Federal District Court Judges according to judges and attorneys. *Criminology, Criminal Justice, Law & Society*, 19, 50-73. <https://ccjls.scholasticahq.com/article/6337-in-their-words-the-sentencing-decisions-of-federal-district-court-judges-according-to-judges-and-attorneys>
- Hartley, Richard D. (2019). Inter-district differences and extra-legal disparity in federal sentencing: The trees are 'substantially' more important than the forest. In C. C. Spohn, & P. Brennan (Eds.), *Sentencing Policies and Practices in the 21st Century. Handbook on Corrections and Sentencing, Vol. 4* (80-113). New York: Routledge.
- Hartley, R. D., Madden, S., & Spohn, C. (2007). Prosecutorial discretion: An examination of substantial assistance departures in federal crack-cocaine and powder-cocaine cases. *Justice Quarterly*, 24, 383-407. doi:10.1080/07418820701485379
- Hartley, R. D., & Tillyer, R. (2012). Defending the homeland: Judicial sentencing practices for federal immigration offenses. *Justice Quarterly*, 29, 76-104. doi: [10.1080/07418825.2011.585140](https://doi.org/10.1080/07418825.2011.585140)
- Johnson, B. D. (2006). The multilevel context of criminal sentencing: Integrating judge and county level influences in the study of courtroom decision making. *Criminology*, 44, 259-298. doi: 10.1111/j.1745-9125.2006.00049
- Johnson, B. D. (2014). Judges on trial: A reexamination of judicial race and gender effects across modes of conviction. *Criminal Justice Policy Review*, 25, 159-184. doi: [10.1177/0887403412463933](https://doi.org/10.1177/0887403412463933)
- Johnson, B. D., & Betsinger, S. (2009). Punishing the model minority: Asian American stereotypes and disparities in criminal punishments. *Criminology*, 47, 1045-1090. doi: 10.1111/j.1745-9125.2009.00169
- Johnson, B. D., Ulmer, J. T., & Kramer, J. H. (2008). The social context of guideline circumvention: The case of federal district courts. *Criminology*, 46, 711-783 doi: 10.1111/j.1745-9125.2008.00125
- Kastellec, J. P. (2013). Racial diversity and judicial influence on appellate courts. *American Journal of Political Science*, 57, 167-183. doi: 10.1111/j.1540-5907.2012.00618
- Kautt, P. (2002). Location, location, location: Interdistrict and intercircuit variation in sentencing outcomes for federal drug-trafficking offenses. *Justice Quarterly*, 19, 633-671. doi: [10.1080/07418820200095381](https://doi.org/10.1080/07418820200095381)
- Kim, B., Spohn, C., & Hedberg, E. C. (2015). Federal sentencing as a complex collaborative process: Judges, prosecutors, judge-prosecutor dyads, and disparity in sentencing. *Criminology*, 53, 597-623. doi: 10.1111/1745-9125.12090
- Kimbrough v. United States, 552 U.S. 85 (2007)
- King, R. D., Johnson, K. R., & McGeever, K. (2010). Demography of the legal profession and racial

- disparities in sentencing. *Law & Society Review*, 44, 1–32. doi: 10.1111/j.1540-5893.2010.00394
- Kulik, C. T., Perry, E. L., & Pepper, M. B. (2003). Here comes the judge: The influence of judge personal characteristics on federal sexual harassment case outcomes. *Law and Human Behavior*, 27, 69–86. doi: 10.1023/A:1021678912133
- Kutateladze, B., Andiloro, N., Johnson, B., & Spohn, C. (2014). Cumulative disadvantage: Examining racial and ethnic disparity in prosecution and sentencing. *Criminology*, 52, 514–551. doi: 10.1111/1745-9125.12047
- Kramer, J. H., & Ulmer, J. T. (2002). Downward departures for serious violent offenders: Local court “corrections” to Pennsylvania sentencing guidelines. *Criminology*, 40, 897–933. doi: 10.1111/j.1745-9125.2002.tb00977
- Kramer, J. H., & Ulmer, J. T. (2009). *Sentencing guidelines: Lessons from Pennsylvania*. Boulder, CO: Lynne Rienner Publishers.
- Lynch, M. (2018). Prosecutorial discretion, drug case selection, and inequality in federal court. *Justice Quarterly*, 35, 1309–1336.
- Metcalf, C. (2016). The role of courtroom workgroups in felony case dispositions: An analysis of workgroup familiarity and similarity. *Law & Society Review*, 50, 637–673. doi: 10.1111/lasr.12217
- Muhlhausen, D. B. (2004). *The determinants of sentencing in Pennsylvania: Do the characteristics of judges matter?* Retrieved from Heritage Foundation website: <https://www.heritage.org/crime-and-justice/report/the-determinants-sentencing-pennsylvania-do-the-characteristics-judges>
- Nardulli, P. F., Fleming, R. B., & Eisenstein, J. (1988). *The tenor of justice: Criminal courts and the guilty plea process*. Urbana, IL: University of Illinois Press.
- Ortiz, N. R., & Spohn, C. (2014). Mitigating the effect of a criminal record at sentencing local life circumstances and substantial assistance departures among recidivists in federal court. *Criminal Justice Policy Review*, 25, 3–28. doi: 10.1177/0887403412459542
- Rodriguez F. S., Curry T. R., & Lee, G. (2006). Gender differences in criminal sentencing: Do effects vary across violent, property, and drug Offenses? *Social Science Quarterly*, 87, 318–339. doi: 10.1111/j.1540-6237.2006.00383
- Savelsberg, J. J. (1992). Law that does not fit society: Sentencing guidelines as a Neo-Classical reaction to the dilemmas of substantivized law. *American Journal of Sociology*, 97, 1346–1381. doi: 10.1086/229904
- The Sentencing Project. (2005). *Racial disparity in sentencing: A review of the literature*. Retrieved from Open Society Foundation website: <https://www.opensocietyfoundations.org/uploads/764bfl50-13d8-4330-b08b-b04ae313308f/disparity.pdf>
- Spohn, C. (1990a). Decision making in sexual assault cases: Do black and female judges make a difference? *Women & Criminal Justice*, 2, 83–105. doi:10.1300/J012v02n01_06
- Spohn, C. (1990b). The sentencing decisions of Black and White judges: Expected and unexpected similarities. *Law & Society Review*, 24, 1197–1216. doi: 10.2307/3053666
- Spohn, C. (2000). *Thirty years of sentencing reform: The quest for a racially neutral sentencing process*. Retrieved from National Criminal Justice Reference Service website: https://www.ncjrs.gov/criminal_justice2000/vol_3/03i.pdf
- Spohn, C. (2009). *How do judges decide: The search for fairness and justice in punishment*. Thousand Oaks, CA: Sage.
- Spohn, C. (2013). The effects of race and ethnicity on federal sentencing outcomes in the guidelines era. *Law and Contemporary Problems*, 76, 75–104. <https://scholarship.law.duke.edu/lcp/vol76/iss1/5>
- Spohn, C., & Fornango, R. (2009). U.S. attorneys and substantial assistance departures: Testing for interprosecutor disparity. *Criminology*, 47, 813–842. doi: 10.1111/j.1745-9125.2009.00163
- Steffensmeier, D., & Britt, C. L. (2001). Judges' race and judicial decision making: Do Black judges sentence differently? *Social Science Quarterly*, 82, 749–764. doi: 10.1111/0038-4941.00057
- Steffensmeier, D., Ulmer, J. T., & Kramer, J. H. (1998). The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black, and male. *Criminology*, 36, 763–797. doi:10.1111/j.1745-9125.1998.tb01265
- Tillyer, R., & Hartley, R. D. (2016). The use and impact of fast-track departures: Exploring prosecutorial and judicial discretion in federal immigration cases. *Crime & Delinquency*, 62(12), 1624–1647. doi: 10.1177/0011128713505481

- Tillyer, R., Hartley, R. D., & Ward, J. T. (2015). Differential treatment of female defendants: Does criminal history moderate the effect of gender on sentence length in federal narcotics cases? *Criminal Justice and Behavior*, 42, 1624–1647. doi: 10.1177/0093854814560624
- Ulmer, J. T. (1997). A processual order approach to studying sentencing guidelines: Contexts, activities, and consequences. *Applied Behavioral Science Review*, 5, 81–100. doi: 10.1016/S1068-8595(97)80008-6
- Ulmer, J. T. (2005). The localized uses of Federal Sentencing Guidelines in four U.S. District Courts: Evidence of processual order. *Symbolic Interaction*, 28, 255–279. doi: 10.1525/si.2005.28.2.255
- Ulmer, J. T. (2012). Recent developments and new directions in sentencing research. *Justice Quarterly*, 29, 1–40. doi: 10.1080/07418825.2011.624115
- Ulmer, J. T., & Bradley, M. (2006). Variation in trial penalties among serious violent offenses. *Criminology*, 44, 631–670. doi: 10.1111/j.1745-9125.2006.00059
- Ulmer, J. T., Eisenstein, J., & Johnson, B. D. (2010). Trial penalties in federal sentencing: Extra-guidelines factors and district variation. *Justice Quarterly*, 27, 560–592. doi: 10.1080/07418820902998063
- Ulmer, J. T., & Johnson, B. D. (2004). Sentencing in context: A multilevel analysis. *Criminology*, 42, 132–172. doi: 10.1111/j.1745-9125.2004.tb00516
- Ulmer, J. T., Painter-Davis, N., & Tinik, L. (2016). Disproportional imprisonment of Black and Hispanic males: Sentencing discretion, processing outcomes, and policy structures. *Justice Quarterly*, 33(4), 642–681. doi: [10.1080/07418825.2014.958186](https://doi.org/10.1080/07418825.2014.958186)
- United States Sentencing Commission. (2010). Demographic differences in federal sentencing practices: An update of the Booker report's multivariate regression analysis. United States Sentencing Commission, Washington, DC.
- United States v. Booker, 543 U.S. 220 (2005)
- Waldfoegel, J. (1991). Aggregate inter-judge disparity in federal sentencing: Evidence from three districts. *Federal Sentencing Reporter*, 4, 151. doi: 10.2307/20639431
- Washington, L. (1994). *Black judges on justice: Perspectives from the bench*. New York, NY: The New Press.
- Wu, J., & Spohn, C. (2010). Interdistrict disparity in sentencing in three US district courts. *Crime & Delinquency*, 56, 290–322. doi: 10.1177/0011128707309997
- Wooldredge, J. (2010). Judges' unequal contributions to extralegal disparities in imprisonment. *Criminology*, 48, 539–567. doi: 10.1111/j.1745-9125.2010.00195
- Zatz, M. (2000). The convergence of race, ethnicity, gender, and class on court decision-making: Looking toward the 21st century. Retrieved from National Criminal Justice Reference Service website: https://www.ncjrs.gov/criminal_justice2000/vol_3/03j.pdf Justice.

About the Authors

Richard D. Hartley Richard D. Hartley is a Professor in the Department of Criminology & Criminal Justice at the University of Texas at San Antonio. Most of his published research examines decision making in criminal court case processing. He is the recent recipient of two grants funded by the National Institute of Justice; a multi-site evaluation of veterans treatment courts, and a project to examine the correlates of prosecutorial decision-making in the federal criminal justice system. His current published research appears in *Justice Quarterly*, *Journal of Criminal Justice*, *Criminal Justice & Behavior*, and *Criminal Justice Policy Review*.

Rob Tillyer Rob Tillyer, Ph.D. is the Associate Dean of Graduate Studies & Research in the College of Public Policy and Professor in the Department of Criminology & Criminal Justice at the University of Texas at San Antonio. His research interests include decision-making within the criminal justice system, situational factors associated with crime, victimology, and crime prevention. Recent journal articles have appeared in *Journal of Research in Crime & Delinquency*, *Justice Quarterly*, *Criminal Justice & Behavior*, and *Crime & Delinquency*. He has also received funding from federal, state, and local governments to oversee several criminal justice related research projects.