THE MULTILEVEL CONTEXT OF CRIMINAL SENTENCING: INTEGRATING JUDGE- AND COUNTY-LEVEL INFLUENCES[*]

BRIAN D. JOHNSON
  University of Maryland
  Department of Criminology and Criminal Justice

KEYWORDS: criminal sentencing, courtroom decision making, contextual disparities, judge and court characteristics, hierarchical linear modeling

This study extends recent inquiries of contextual effects in sentencing by jointly examining the influence of judge and courtroom social contexts. It combines two recent years of individual sentencing data from the Pennsylvania Commission on Sentencing (PCS) with data on judicial background characteristics and county court social contexts. Three-level hierarchical models are estimated to investigate the influence of judge and county contexts on individual variations in sentencing. Results indicate that nontrivial sentencing variations are associated with both individual judge characteristics and county court contexts. Judicial background factors also condition the influence of individual offender characteristics in important ways. These and other findings are discussed in relation to contemporary theoretical perspectives on courtroom decision making that highlight the importance of both judge and court contexts in sentencing. The study concludes with suggestions for future research on contextual disparities in criminal sentencing.

As a fundamental mechanism of social control in society, criminal sentencing is a forum where broad sociological concerns, such as equality under the law, meet individualized decision making constraints, such as locally varying courtroom norms and individual courtroom actor influences. Historical interest in this area has long been driven by philosophical and political concerns surrounding discrimination in the criminal justice system (Sellin, 1928), with modern research largely focusing on individual-level disparities among similarly situated offenders (for example, Albonetti, 1997; Bushway and Piehl, 2001; Steffensmeier, Ulmer, and Kramer, 1998). More recently, though, research on criminal sentencing has emphasized the importance of considering contextual influences, such as the characteristics of individual decision makers (for example, Steffensmeier and Britt, 2001) or the role of courtroom social contexts (for example, Johnson, 2005; Ulmer and Johnson, 2004). Despite the increased recognition of these important influences, empirical work remains underdeveloped. Virtually all research has been restricted to separate and isolated investigations of judge- or county-level influences, limiting our understanding of the ways that individual sentencing outcomes vary across both judges and county courts. This study therefore integrates judge- and county-levels of analysis to more accurately specify the complex relationships among courtroom actor background characteristics, courtroom-level social contexts, and individual-level disparities in sentencing.

From a societal welfare perspective, the potential for contextual variations in sentencing raises important issues of unwarranted disparities in criminal justice as well as larger issues of inequality in society. From a methodological standpoint, studies that fail to incorporate either judge or county courtroom influences may risk omitted variable bias and model misspecification. These considerations may be especially important given the recent proliferation of sentencing guidelines designed to limit judicial discretion and reduce or eliminate variations in sentencing outcomes across judges and jurisdictions. With few exceptions, empirical work has not investigated the extent to which guidelines have resulted in sentencing equivalencies across judges and courts. Although research on courtroom contexts has advanced notably in recent years (Britt, 2000; Johnson, 2005; Kautt, 2002; Ulmer and Johnson, 2004), studies of judge-level factors have lagged in methodological sophistication. The goal here is to advance research on contextual effects in sentencing by incorporating a broad array of theoretically important judge- and county-level influences while addressing a number of limitations of earlier work on judge-level effects in criminal sentencing.

First, this work investigates a broader range of contextual variables than most earlier studies. Research on judge effects in particular has

incorporated a limited number of variables (Walker and Barrow, 1985), often focusing on a single judicial characteristic such as race (for example, Spohn, 1990a), ethnicity (Holmes, Hosch, and Daudistel, 1993), or gender (Gruhl, Spohn, and Welch, 1981). This work includes both more extensive controls at the judicial level and additional variables of theoretical interest, such as caseloads and measures of courtroom efficiency that better address emergent concerns about the organizational climate of courtroom decision making.

Second, virtually all studies of contextual effects in sentencing to date have been limited to a single level of contextual analysis (for an important exception, see Myers and Talarico, 1987). Given that judges, like other organizational actors, are embedded in interpersonal social networks that shape their behavioral expectations and cultural norms (Granovetter, 1985), the absence of controls for the organizational context of the court is an important limitation. Similarly, research on courtroom contexts has failed to fully consider the role of the individual judge. Although judges share similar organizational contexts, they remain individual actors, subject to influences from idiosyncratic backgrounds and personal experiences (Hogarth, 1971). By incorporating both judge- and county-level contextual factors, then, this study furthers our understanding of the relative importance of judge and county court contexts in criminal sentencing decision-making processes.

Finally, it also offers the first analysis of judicial sentencing factors to properly account for the hierarchical nesting of cases within sentencing judges (as well as judges within courts), providing more reliable estimates of judge- and county-level effects in criminal sentencing. No extant research on judicial effects has yet accounted for this structuring. Although appropriate methodologies have been increasingly adopted and shed new light on the study of county courtroom factors (for example, Bontrager, Bales, and Chiricos, 2005; Britt, 2000; Helms and Jacobs, 2002; Johnson, 2005; Kautt, 2002; Ulmer and Johnson, 2004), the same is not true for studies of judge-level effects in sentencing.

JUDGE- AND COUNTY-LEVEL CONTEXTS

The importance of social contexts in criminal sentencing was perhaps best enumerated in the seminal work of Myers and Talarico (1987), which analyzed a variety of judge-, court-, and county-level predictors of sentencing in Georgia. They argued that criminal sentencing decisions varied along three dimensions: county, court (including judge factors), and temporal contexts. Their findings, though complex, suggested a number of modest direct effects from both judicial background characteristics (including age, religion, and prosecutorial experience) and county

contextual factors (including urbanization, socioeconomic conditions, and racial composition). Although they concluded that individual offender and case characteristics outweigh the influence of social contexts, they did note that both judge- and county-level predictors exert important conditioning effects on individual criminal sentencing outcomes. This led them to conclude that "in contrast to their limited direct effects, [contextual] variables operate as impressive conditioners of differential treatment" in sentencing (Myers and Talarico, 1987: 122).

Although this study established the fundamental importance of examining contextual effects in sentencing, appropriate statistical techniques for analyzing multilevel data were not yet commonplace, and several key judicial characteristics lacked variability. For instance, Georgia at the time had no minority judges and almost no female judges sentencing criminal cases. Myers and Talarico realized this limitation: "As a result of this homogeneity, we could not reliably estimate the effects of several attributes of concern to other researchers" (1987: 29). This is the only existing study to simultaneously examine measures of both judge- and county-level social contexts. Related research has focused instead on judicial background characteristics or county courtroom factors in isolation.

JUDICIAL CONTEXT

Long ago Hogarth averred, "one can explain more about sentencing by knowing a few things about a judge than by knowing a great deal about the facts of the case" (1971: 350). Despite this provocative argument, few empirical studies have attempted to assess the overall degree to which sentencing outcomes vary across judges. Most research on inter-judge disparity has focused on the influence of judicial background characteristics. Collectively, these studies have reported few substantial findings, leading a number of scholars to conclude that the cumulative effects of sentencing guidelines, legal training, and judicial socialization processes have resulted in the emergence of a largely homogenous judiciary, deciding similar cases in similar ways (for a recent review see Zatz, 2000: 509). However, this conclusion may be hasty, given the relative paucity of research devoted to judicial variations in sentencing. Whereas a copious literature has developed around individual offender attributes, existing research on judicial characteristics is limited in both scope and number (Zatz, 2000).

Many studies of judge-level effects have examined courtroom decisions in one or two particular cities (Gruhl, Spohn, and Welch, 1981; Kritzer and Uhlman, 1977; Uhlman, 1978), for particular types of cases in federal district courts (Kritzer, 1978; Walker and Barrow, 1985), in state and

federal supreme courts (Songer and Crews-Meyer, 2000), or for specific types of cases within state-level trial courts (Spohn, 1990a, 1990b). Overall, though, few studies have systematically examined the impact of judge characteristics in criminal courts for a large number of judges or for a variety of different cases (but see Myers and Talarico, 1987; Steffensmeier and Hebert, 1999).

Moreover, the limited number of studies that have focused on judge-level variables in criminal sentencing have often relied on small samples (for example, Kritzer and Uhlman, 1977), been limited to a single factor (for example, Gruhl, Spohn, and Welch, 1981), or used methodological techniques that fail to account for the nested nature of sentencing data (for example, Myers and Talarico, 1987; Steffensmeier and Hebert, 1999). Most that have investigated judicial background factors in criminal sentencing also rely on now-outdated datasets. For instance, four separate analyses use data from "Metro City" that dates to 1968 (Gruhl, Spohn, and Welch, 1981; Kritzer and Uhlman, 1977; Ulhman, 1978; Welch, Combs, and Gruhl, 1988). Even Steffensmeier and Britt's (2001) recent study uses data from the early 1990s that is more than 10 years old. Given the ever-changing nature of sentencing guidelines and the limited scope of existing research, modern investigations of judicial background influences offer an important contribution to current research on criminal sentencing.

Early work often focused on a single characteristic, such as the gender or race of the sentencing judge. Both the Kritzer and Uhlman (1977) and the Gruhl, Spohn, and Welch (1981) studies of judge's gender in "Metro City" identified few differences between male and female judges, and Uhlman's (1978) suggested that black judges sentence black offenders more leniently than white judges do. In line with this early work, Spohn (1990a, 1990b) examined race and gender effects in Detroit Recorder's Court. In her first study, she examined violent felonies and found that black judges were significantly less likely than their white counterparts to sentence offenders to prison, though both black and white judges sentenced black offenders more harshly. Still, her summary conclusion was that judicial race had relatively little predictive power. In the second study, she concluded that there were few differences between male and female judges in sentencing sexual assault cases, though she did report that female judges sentenced offenders to almost 4 years of additional incarceration— a disparity that was even greater when black female judges were considered separately.

More recent work has incorporated additional judicial factors, such as the age of the judge, tenure on the bench, marital status, and experience as a prosecutor (Frazier and Bock, 1982; Myers and Talarico, 1987; Steffensmeier and Hebert, 1999; Welch, Combs, and Gruhl, 1988). These studies too have uncovered few strong influences from judicial

background factors. The two most recent and comprehensive studies reexamined gender and race effects in select Pennsylvania counties. In the first, Steffensmeier and Hebert (1999) reported that after controlling for several judge-level factors, female judges sentenced more harshly than male judges, especially when faced with repeat black offenders. Similarly, in the most recent treatment of race effects, Steffensmeier and Britt (2001) reported that, contrary to earlier research, black male judges in the four counties they examined were overall more likely to incarcerate criminal defendants, both black and white.

Overall, work on judge characteristics suggests they often exert significant but modest influences, though inconsistencies among studies are common. In part, this may reflect the fact that few large-scale systematic studies of judge effects in sentencing have been undertaken, limiting what we know about the extent to which sentencing outcomes vary across judges, are associated with particular judicial background factors, or are influenced by judicial case processing characteristics. On the other hand, considerably more research has been devoted to understanding the contextual influences of courtroom and community characteristics in criminal sentencing.

COURTROOM CONTEXT

Research on contextual effects in sentencing demonstrates the importance of examining a number of different courtroom and community characteristics. Studies conducted at both the state level (for example, Britt, 2000; Dixon, 1995; Ulmer and Johnson, 2004) and federal level (for example, Albonetti, 1997; Kautt, 2002; Mustard, 2001) suggest that individual outcomes vary across court and community contexts. A number of studies have focused on the influence of particular courtroom characteristics, such as courtroom caseloads or court resources (for example, D'Alessio and Stolzenberg, 1997; Dixon, 1995; Kautt, 2002), whereas other studies have targeted community characteristics, such as the county-level crime rate, or the political or racial composition of the community (for example, Bontrager, Bales, and Chiricos, 2005; Crawford, Chiricos, and Kleck, 1998; Helms and Jacobs, 2002). Relatively few studies, though, have incorporated a broad spectrum of both courtroom and community influences on sentencing (but see Myers and Talarico, 1987; Johnson, 2005; Ulmer and Johnson, 2004).

Although a number of studies report significant influences from courtroom- and community-level social contexts, these findings are often difficult to summarize concisely. Partly this is the result of the subtle and often indirect nature of contextual effects, and partly it reflects the fact that prior research often relied on regression techniques that overstate the

statistical significance of contextual influences (Ulmer and Johnson, 2004). Several recent studies, however, have used more appropriate multilevel models to correct for this bias in their investigations of courtroom sentencing contexts.

Britt's (2000) examination of racial disparity in Pennsylvania found mixed support for the role of urbanization and racial composition on incarceration and sentence lengths, and little support for the role of income or crime rates on sentencing. He reported that, contrary to early work (Myers and Talarico, 1987), these measures of county context did little to condition the individual effect of race. Kautt's (2002) analysis of federal drug trafficking offenses similarly found mixed evidence for contextual influences. Specifically, she found no discernable effects for district-level minority composition, unemployment rate, population size, or drug caseload factors. However, she did find that the departure rate of the district was significantly related to individual offender sentence lengths, offering some support that individual sentences are embedded in and conditioned by aggregate, organizational patterns of courtroom actor behavior.

More recently, Ulmer and Johnson (2004) reported a number of significant influences on the judicial decision regarding incarceration, though these factors mattered less for length of imprisonment. Offenders sentenced in large courts were about half as likely to be incarcerated as those in medium or small courts, and both the caseload pressure of the court and the availability of local jail space were significantly related to incarceration in predicted directions. No evidence was found in this study, though, for the importance of community-level factors such as the socioeconomic condition, racial composition, political makeup, or crime rate of the surrounding community.

Finally, two recent studies have extended this line of inquiry to additional outcomes of interest. Bontrager and her colleagues (2005) examined the importance of court contexts for the decision to withhold adjudication of guilt in a sample of Florida felons. Using a racial threat perspective, they examined racial and ethnic disparities and concluded that the effects of race and ethnicity varied significantly by percentage minority and level of concentrated disadvantage. Johnson (2005) examined variation across courts in circumventing guidelines and found that the likelihood of receiving a departure from the sentencing guidelines also varied significantly across courts. In particular, offenders processed in large courts were about twice as likely to be sentenced below the guidelines and half as likely to be sentenced above them. The likelihood of departure was also influenced by the departure rate of the court, its caseload pressure and trial rate, and the percentage Hispanic in the

community. These various measures of social context significantly moderated the influence of several individual sentencing considerations.

Although a useful foundation, none of these studies consider the potentially important influences of judge-level factors in sentencing. To the extent that these influences are tied to courtroom community influences, research that focuses on only one or the other may be missing important elements in the courtroom decision-making process. For instance, research on judge effects has suggested that minority judges may be less punitive (Spohn, 1990a), and research on the size of the court suggests that large urban courts may sentence less severely (Ulmer and Johnson, 2004). To the extent that minority judges congregate in large urban courts, however, these two effects are likely to be confounded. To better parcel out the relative contribution of judge and county court contexts, then, researchers must integrate these levels of analysis in the study of criminal sentencing. To accomplish this, I apply a number of contemporary theoretical perspectives on criminal sentencing to suggest various hypotheses regarding the influence of judge- and county-level predictors in the courtroom decision-making process.

THEORETICAL IMPORT

Contemporary theoretical sentencing perspectives emphasize the fact that courtroom decision making is not only influenced by individual case and offender characteristics, but is also a product of more subtle influences associated with the courtroom actors themselves and their particular social environments (Eisenstein, Flemming, and Narduli, 1988; Gibson, 1983; Ulmer, 1997). Even under presumptive sentencing guidelines, courtroom decision making largely remains a substantive process (Savelsberg, 1992). As Albonetti argued, judicial sentencing decisions are often constrained by a lack of complete information regarding an offender's dangerousness and likelihood of future offending, which leads courtroom actors "to reduce uncertainty by relying upon a rationality that is the product of habit and social structure" (1991: 249). Under this condition of bounded rationality, sentencing outcomes result not only from the formal considerations embodied in sentencing guidelines, but also from the substantive concerns associated with courtroom actor interpretations of individual offender traits that are tied to fundamental courtroom considerations.

According to Steffensmeier and his colleagues, these considerations revolve around three primary "focal concerns" of sentencing: offender culpability, community protection (that is, offender dangerousness), and practical constraints surrounding individual offenders and organizational resources (Johnson, 2003; Steffensmeier and Demuth, 2000; Steffensmeier,

Ulmer, and Kramer, 1998). Because judges and other courtroom actors are forced to make decisions under time and information constraints, they are likely to use stereotypical attributions—or patterned responses (Hawkins, 1981)—that relate their interpretations of different focal concerns to particular offender and case characteristics. This process allows for the subtle influences of experiences, prejudices, and stereotypes, as well as idiosyncratic interpretations by different judges, to enter into the courtroom decision-making process. To the extent that individual judges have unique attitudes, beliefs, and background experiences, their situational interpretations of the relative import of different focal concerns is likely to differ.

Gibson, for instance, has argued that "there can be little doubt that the behavior of judges is in fact predictable from their backgrounds" (1983: 23). Although these individual differences may be mitigated by the collective adoption of a common "subculture of justice" (Frazier and Bock, 1982), encouraged through judicial selection procedures (Levin, 1977), socialization experiences (Spohn 1990a; Steffensmeier and Hebert, 1999), and organizational norms of conformity (Eisenstein, Flemming, and Nardulli, 1988), the inherent uncertainty in courtroom decision-making processes still presents important opportunities for variations in judicial sentencing to occur. These variations, then, are theoretically expected to be tied to the background and case processing experiences of the sentencing judge. As Hogarth aptly put it, judges

> differ widely in their purposes, their views as to the effectiveness of different kinds of sentences, the criteria applied in deciding between different kinds of sentences, the ways in which conflict between the offender's needs and community protection is resolved, the amount of information support they have for their views, and the kinds of situation in which they experience difficulty in sentencing. (1971: 91)

This suggests persuasive theoretical arguments for expecting interjudge variation in sentencing even under presumptive sentencing guidelines. This study further explores the possibility by explicitly estimating variations in judicial sentencing patterns before attempting to explain them using a variety of judge-level predictors of sentencing.

Contemporary theoretical perspectives also further highlight the importance of courtroom social contexts. As Albonetti's (1991) work suggests, uncertainty in organizational decision making is likely to lead to the emergence of "organizational arrangements" such as "going rates" (Eisenstein, Flemming, and Nardulli, 1988), which are routinely dispensed for "normal crimes" (Sudnow, 1965) to make the decision-making process more predictable. The development of these organizational arrangements, however, varies across courts. In line with courtroom community theory,

which views criminal courts as unique communities, or social worlds (Ulmer, 1997), sentencing outcomes are decided in accordance with shared workgroup expectations that emerge in the context of locally varying courtroom normative environments (Dixon, 1995; Eisenstein, Flemming, and Nardulli, 1988; Ulmer and Johnson, 2004). Such factors as the location and size of the court, its level of bureaucracy, and the case processing strategies that develop all foster the emergence of a unique court culture, resulting in locally varying sentencing norms that perpetuate sentencing variations across courts.

The focal concerns perspective also highlights the fact that judges and other courtroom actors are sensitive to organizational resources and constraints as well as to community pressures (Steffensmeier, Ulmer, and Kramer, 1998). Because the availability of courtroom resources and the salience of different organizational concerns vary across counties, these differences further exacerbate between-court variations in sentencing. Moreover, courts are embedded in community environments that also differ in important ways; for instance, social and racial threat processes (Blalock, 1967; Liska, 1992) may be related to differences in economic prosperity or minority presence across communities. According to modern theoretical perspectives on sentencing, then, differences among sentencing judges and differences among courtroom social environments are both likely to result in significant variations in individual sentencing outcomes. Moreover, the relative weight placed on individual sentencing factors is also expected to vary across judges and county courts (Gibson, 1983; Myers and Talarico, 1987; Ulmer, 1997). Hence this study investigates the following hypotheses.

1. Sentence severity will vary significantly across both judges and county courts.
2. The effects of individual-level sentencing factors will also vary across judges and courts.

If sentencing outcomes do vary significantly across judges and courts, then the next task is to explain these variations with judge- and county-level predictors of sentencing. Although a number of judge-level predictors have been hypothesized, by far the most attention has been devoted to judicial demographic characteristics such as the race and gender of the judge. Although research regarding these effects has been decidedly mixed (for example, Spohn, 1990a; Steffensmeier and Britt, 2001), theoretical arguments point to leniency among female and minority judges. Long ago, Goldman argued that "a judge who is a member of a racial minority or a woman cannot help but bring to the bench a certain sensitivity—indeed, certain qualities of the heart and mind—that may be particularly helpful in dealing with racial and sexual discrimination issues" (1979: 494). More generally, Welch, Combs, and Gruhl (1988: 127) have

argued that black judges tend to be more liberal and therefore may be "more sympathetic to criminal defendants than white judges" because "liberal views are associated with support for the underdog and the poor." The same may also hold true for the gender of the sentencing judge. Gruhl and colleagues (1981), for instance, have argued that women tend to be more liberal than men on a host of social issues, which may result in more lenient sentencing patterns. To the extent that the race or gender of the judge is tied to early socialization processes (that is, race and gender roles), as well as differences in attitudes and background experiences, one would expect systematic differences in sentencing outcomes to emerge. Specifically, one might expect both minority and female judges to be more receptive to lower class interests and have more liberal punitive philosophies, thus favoring more lenient outcomes, as suggested in the following hypotheses.

3. Minority judges will sentence more leniently than white judges.
4. Female judges will sentence more leniently than male judges.

Although research to date has not examined them, judicial caseload factors are also an important element in the judicial decision-making process. As Hogarth (1971: 217) expressed it, "the work-load of a magistrate has a direct influence on the way in which he makes decisions." Not only does a judge's caseload affect the amount of time he has to consider a given case, but it also influences the "degree to which he is able to devote time and thought to general considerations in sentencing" (Hogarth, 1971: 217). Courtroom actors are keenly aware of organizational maintenance goals, such as the efficient disposal of caseloads. Judicial caseload pressure is therefore likely to influence sentencing patterns across judges. Given a heavy workload, it may be advantageous for judges to sentence more leniently in order to expedite case processing and conserve courtroom resources (Dixon, 1995). In addition, the types of cases a judge hears may also influence his or her decision making. As Emerson (1983: 426) persuasively argued, "the makeup of the overall 'stream of cases'. . . provides a background against which the classification of particular cases in organizationally relevant ways will be made." Individual case considerations are determined in relation to the judge's case flow. Judicial assessments of case seriousness are at least partially the product of the overall seriousness of other cases the judge has experienced. A judge who sentences a high proportion of violent cases, for instance, is likely to establish a higher threshold for evaluating serious crime, which may result in less overall sentencing severity. Two additional hypotheses are therefore predicted.

5. Judges with heavier caseloads will sentence more leniently.
6. Judges with heavier violent caseloads will sentence more leniently.

Extant theory also suggests that a number of courtroom characteristics are important predictors of individual sentencing outcomes. Although important differences are likely to exist in the ways that different judges sentence similar cases, "unequal sentences for similar offenses may also result from differences in the social contexts in which the courts operate, such as differences in the crime rate, or in public opinion, or in the resources to deal with offenders available locally" (Hogarth, 1971: 7). As Myers and Talarico (1987) pointed out, "judges do not function in isolation. Their preferences and expectations are substantially conditioned by the environment in which they work" (7). Judicial decision making is shaped and constrained by the organizational, political, and social milieu surrounding the court. Judges are acutely aware of caseload management demands, local political pressures, and the social consequences of their decisions in the community. Therefore the various dynamics that contribute to the courtroom community and its surrounding social environment are likely to influence sentencing outcomes across contexts.

Specifically, courtroom community theory argues that the size of the court is one of the most crucial factors to understanding contextual variations in sentencing. This expectation has received support in recent empirical work (for example, Johnson, 2005; Ulmer and Johnson, 2004). The size of the jurisdiction is tied to variations in important organizational and cultural features of the court community, such as its autonomy from external controls, its public visibility regarding routine cases, and the diversity of different criminal cases sentenced (Eisenstein, Flemming, and Nardulli, 1988). These influences are hypothesized to lead to relative severity in small jurisdictions and relative leniency in large jurisdictions. The following relationship is therefore expected.

7. Small courts will sentence more severely than large courts.

Locally varying criminal justice resources and community pressures are also likely to influence general patterns of sentencing severity. From a focal concerns perspective, judges should be sensitive to organizational constraints and local environmental pressures. Such factors as the overall departure and trial rates (Kautt, 2002; Johnson, 2005) and the availability of local jail capacity (D'Alessio and Stolzenberg, 1997; Ulmer and Johnson, 2004) have been shown to affect sentencing decisions. If, as courtroom community theory suggests, different courts develop different standards on the appropriate balance between formal guideline recommendations and established courtroom sentencing norms, then judicial sentencing decisions should be influenced by the departure rate of the court. Higher departure rates may reflect a larger disjuncture between informal sentencing norms and formal guidelines recommendations, or they may indicate differential embeddedness of substantively rational sentencing concerns across courts (Ulmer and Kramer, 1996). Either way,

lower guidelines compliance likely reflects a more flexible courtroom normative environment, which is likely to result in more lenient sentencing patterns (Kautt, 2002).[1] Moreover, if judges are sensitive to courtroom resources, as courtroom efficiency theory suggests, the increased availability of local jail space should lead to increased judicial use of incarceration. The following courtroom-level hypotheses are therefore also investigated.

   8. Higher guidelines departure rates will result in more lenient
      sentencing outcomes.
   9. Available jail capacity will increase the likelihood of
      incarceration.

   Although the primary focus of the present study is on judge-level effects in sentencing, the fact that previous research on these critical courtroom contexts has not controlled for judge-level factors makes it equally important to replicate these findings.

   There are other theoretical reasons to expect important interactions between judge and offender attributes. This is because judicial attributions of offender culpability and dangerousness are not only influenced by the characteristics of the judge and county court, but also by perceived similarities and differences between the judge and the offender (Hawkins, 1981: 211). White judges, for instance, may be more likely to have negative stereotypical attributions regarding the dangerousness and culpability of minority offenders. Minority judges, on the other hand, may be more sensitive to racial issues, such as the overrepresentation of young minority offenders in prison, which may result in further differences in their treatment of individual offenders. Similar expectations may apply to gender as well. Given paternalistic norms of chivalry in society, male judges should be more likely to be influenced by stereotypical attributions tied to female offender status.[2] Whereas female judges may be more likely to treat male and female offenders as "criminal equals," male judges are likely to view female offenders as less dangerous or less culpable (Gruhl, Spohn, and Welch, 1981). Different sentencing patterns may therefore emerge with respect to different combinations of judge-offender race and gender. In addition, judicial attributions of criminal seriousness are also

---

1. Because similar numbers of cases received downward (N = 6,566) and upward (N = 7,046) departures, this measure is not a simple proxy for sentencing leniency as one reviewer suggested.

2. Though beyond the scope of the present investigation, it is important to note that these judge/offender interactions may be offense specific. For instance, some scholars argue that gender interactions should be more pronounced for "female-sensitive" crimes such as rape or sexual assault (for example, Kulik, Perry, and Pepper, 2003). Arguably, then, the present study offers a conservative test of this hypothesis.

tied to the relative prevalence of different crime types in a judge's caseload (Emerson, 1983). In particular, a judge who sentences a high proportion of violent crimes is likely to view individual violent offenses as more normative, which may color his or her attributions of seriousness for these and other crimes. A judge who sentences very few violent offenses, on the other hand, is likely to perceive these cases as relatively more serious. Given the fundamental importance of race and gender in sentencing, and the potential salience of caseload compositions for judicial determinations of relative case severity, the following interactions between individual- and judge-level characteristics in sentencing are examined:

10. Minority judges will sentence minority offenders more leniently.
11. Male judges will sentence female offenders more leniently.
12. The effect of a violent crime will be greater for judges sentencing less violent crimes.

DATA AND METHODS

To examine these theoretical expectations, this study draws on three separate but related data sources: individual offender data maintained by the Pennsylvania Commission on Sentencing (PCS); judicial background characteristics independently obtained and augmented using judicial biographies of Pennsylvania judges; and county-level courtroom data that combine independently collected data with information from the Administrative Office of Pennsylvania Courts and the U.S. Census Bureau. The individual-level data is comprised of 2 recent years (1999 and 2000) and is limited to cases sentenced under the 1997 Pennsylvania Sentencing Guidelines, and to those cases providing necessary information on judge and county identifiers. The data were also restricted to include only the most serious offense per judicial transaction. These criteria resulted in 148,590 individual-level cases, 79,333 of which resulted in incarceration.[3] The PCS data contain rich detail about individual offense

---

3. In total, 788 cases were removed (0.5 percent of all cases) from the analyses after the most serious offenses were identified. These included cases missing requisite county or judge identifiers, as well as other problematic cases. Seventeen cases were deleted for missing county identifiers, eighty-one for missing judge identifiers, seven for missing both county and judge identifiers, forty-seven for missing information on offense severity, 111 that were murder 1 or murder 2 cases where the guidelines do not apply, thirty-one for missing Prior Record Score, two for erroneous Offense Gravity Score, PRS combination (OGS=4, PRS=8, which is not possible under the 1997 guidelines), and 492 cases sentenced by senior, retired, or traveling judges who sentenced fewer than fifteen cases over the 2-year period.

and offender characteristics, including a wide range of both legally prescribed (that is, legal) and legally proscribed (that is, extralegal) variables, arguably making it one of the best available sources for studying sentencing outcomes at the state level. Despite its strengths, though, these data share weaknesses common to the vast majority of research on individual sentencing outcomes, such as a lack of information on offender-victim relationships, socioeconomic offender status, and information on earlier stages of criminal processing.

The second level of analysis consists of a wide range of background information on the sentencing judges. Cases were limited to judges who sentenced a minimum of fifteen cases over the 2-year period, expressly to eliminate senior or retired judges for whom biographical data were not available, and to remove "traveling" judges (see Levin, 1977) who sentenced a small number of cases across several county courts. After these judges were removed, information remained for 303 sentencing judges.[4]

The third and last level of analysis—the county-level courts—consists of information on both courtroom community characteristics and the surrounding community-level social environment. Although courtroom- and county-level factors may be substantively distinct from one another, they operate at the same level of analysis. The sixty-seven Pennsylvania counties are grouped into the sixty judicial districts presided over by the same judges. These districts, then, serve as the county-level of analysis.[5] Collectively, these interrelated levels of data provide one of the most comprehensive contemporary resources for examining the multilevel contexts of criminal sentencing.

DEPENDENT VARIABLES

Sentencing scholars have argued that the sentencing decision can be broken down into two distinct but related outcomes: the decision to incarcerate and if incarcerated the decision regarding length of sentence (Wheeler, Weisburd, and Bode, 1982). The present work follows this

---

4. A total of 392 judges sentenced criminal cases in Pennsylvania in 1999 and 2000. Limiting the analysis to judges who sentenced a minimum of fifteen cases removed eighty-nine judges, twenty-nine of whom sentenced only one case over the 2-year period. Collectively, the eighty-nine judges sentenced only 492 (.3 percent) of the total cases in the 2 years, and biographical information for most of these judges was not available.

5. The following seven sets of small rural counties share the same judges: Snyder and Union; Colombia and Montour; Forest and Warren; Franklin and Fulton; Perry and Juniata; Wyoming and Sullivan; and Cameron and Elk. Examination of the county-level predictors for these dyads indicated they were virtually identical in their aggregate characteristics.

convention.[6] For the in-out decision, incarceration was coded 1 if the offender was sentenced to any length of confinement in a county jail or state prison and 0 if they were sentenced to any combination of nonincarceration options (for example, probation, restitution).[7] For those incarcerated, then, the sentence-length variable was coded to equal the natural logarithm of the minimum months of incarceration. Because sentence-length data is positively skewed, the error terms in a linear regression tend to be curvilinear, leading to misestimated standard errors and potential estimation bias (Bushway and Piehl, 2001; Kurlychek and Johnson, 2004). The log transformation addresses this issue by normalizing the skewed distribution. It also addresses the fact that additional months of incarceration may have different meaning for different sentence lengths. For instance, a 6-month increase in one's sentence from 6 months to 12 months is more consequential than an equal increase from 60 to 66 months. The former doubles the sentence whereas the latter increases it by only 10 percent. The logged outcome accounts for this by expressing sentence length in terms of the proportional increase in length associated with a unit increase in the explanatory variable of interest.

CONTEXTUAL PREDICTORS OF SENTENCING

Because it is possible that differences among judges and counties may arise from differences in case compositions, it is important to control for various individual case-level sentencing factors when assessing contextual variations in sentencing. Given the broad range of variables included in this analysis, full descriptions are provided in the appendix. Collectively,

---

6. There is some debate regarding the appropriateness of modeling sentencing outcomes in one versus two stages (Albonetti, 1997; Bushway and Piehl, 2001; Kurlychek and Johnson, 2004). Although both conventions offer certain advantages, this study uses the two-stage method for two primary reasons: the hierarchical modeling techniques (HLM) used in the present study do not yet accommodate one-stage estimation procedures such as tobit analysis, and prior research indicates that judge-level factors may exert different influences on the decision to incarcerate versus the decision regarding length of sentence (for example, Spohn, 1990b).

7. Recent scholarship argues for separate examination of jail and prison sentences (Holleran and Spohn, 2004). Multinomial analyses separately comparing probation to jail and prison produced parallel findings for judge- and county-level measures. Differences between the likelihood of jail and the likelihood of prison were uniformly small and statistically insignificant, with the sole exception being for violent crime caseload which had a slightly larger effect on prison (prison: $b = .04$, SE = .01; jail: $b = .01$, SE = .01). These supplemental findings are available by request, but given the similarity of these influences along with the complexity of the present analyses these categories are combined in a single, total incarceration variable here for ease of presentation.

the individual-level factors capture an important mix of legally prescribed factors, including the severity and type of offense, the criminal history of the offender, and the presumptive sentence recommendation, as well as theoretically relevant legally proscribed factors, such as the age, gender, and race-ethnicity of the defendant and the mode of conviction. In addition, because sentence-length models include only those cases that received an incarceration sentence, the "hazard rate" is included as an additional predictor of sentence length to correct for potential selection bias (Berk, 1983; Heckman, 1976).[8] This additional variable was calculated using the Heckman two-step command in Stata 8.2 and then imported into the hierarchical models examining sentence length.

Of primary interest are the judge-level factors, particularly their race, gender, and caseload. Judge race was coded as a dichotomy, with minority judges coded 0 and white judges coded 1. Ideally, it would have been preferable to further distinguish among the minority judges, but this level of detail was not available. Judge gender was also coded as a dichotomous variable, with male judges coded 0 and serving as the reference category, and female judges coded 1. Caseload characteristics were captured with two variables measuring composition and pressure. For composition, three continuous variables identifying the percentage of cases sentenced by the judge for violent, property, and drug cases were coded as percentage violent, percentage property, and percentage drugs. For pressure, the number of cases sentenced by the judge in his or her judicial district was divided by two to represent the average number of criminal cases sentenced in a year. This dividend was subsequently divided by 100 to aid interpretation. Judicial caseload, then, serves as a measure of the relative caseload pressure on the judge at sentencing. Because Pennsylvania judges may adjudicate both criminal and civil cases, though, this is an imperfect measure. Unfortunately, no data on civil caseloads were available. Still,

---

8. There is considerable controversy surrounding the merits and demerits of different approaches to correcting selection bias in criminological research. Some scholars argue for the importance of including Heckman's (1976) correction factor (Peterson and Hagan, 1984) while others maintain that it has the potential to introduce multicollinearity and bias parameter estimates (Stolzenberg and Relles, 1990). Given this concern, the analysis for sentence length was reexamined without the hazard rate. These results produced identical findings for all of the contextual predictors with one exception—the marginally significant influence of large courts on sentence length became insignificant in the uncorrected model. While Heckman's correction appropriately applied can address bias introduced by cases that are convicted but not incarcerated, it is important to note that it does not account for cumulative selection bias from earlier decision-making stages of the system (for example, cases lost from arrest to conviction). This is an important limitation of the present study and is characteristic of much research on criminal sentencing.

given the limited research on judicial caseload effects, examination of criminal caseload pressure remains a worthwhile endeavor.

A broad range of judicial control variables were also incorporated into the analyses, including age and tenure, legal experience, and marital and military status (see appendix). Ideally, political affiliation would have also been examined. However, once elected, Pennsylvania judges run in nonpartisan retention elections (Steffensmeier and Britt, 2001), which makes it difficult to identify their political parties, particularly for those who have sat on the bench for multiple terms, as is often the case.

A variety of county- and court-level contextual measures are also included in the analyses. Of particular interest in the present work are the influences of court size, guidelines departure rate, and local jail capacity. Following Ulmer (1997), courtroom size was trichotomized into large, medium, and small based on both the number of trial judges and the proportion of cases sentenced in each county. The departure rate was calculated as the percentage of cases sentenced outside of the presumptive guidelines range. This variable is used to examine potential variations in the embeddedness of informal sentencing patterns and formal guideline recommendations across courts (Kautt, 2002). The trial rate, measured as the percentage of cases convicted by jury, is also included to control for case processing differences across jurisdictions. Finally, jail space was calculated as the total number of jail beds in each county divided by the number of cases in that county. Therefore, the higher the ratio of jail beds to cases, the higher the relative jail capacity. Several measures of the surrounding social environment of the court were also included as control variables. These included the percentage Hispanic, percentage unemployed, and percentage Republican in the community. Collectively, these variables represent controls for social, economic, and political differences across county-level courts.[9]

---

9. Additional county-level courtroom control variables were also examined, such as the crime rate and percentage black in the county. However, because these additional variables were highly correlated with other theoretically important contextual factors and because they did not improve model fit, they were removed from the final analyses. Multiple model specifications were compared to ensure that the significant results in the reported specification were not the result of these omitted variables. Composite measures combining related items were avoided on theoretical grounds. As one example, percentage black and court size were highly correlated but they clearly represented distinct theoretical processes (that is, racial threat versus court communities). Model comparisons demonstrated that court size was the stronger and more robust predictor of criminal sentencing so it was retained in the model. To tap into racial threat, percentage Hispanic was used in place of percentage black, which still offers theoretical leverage on racial threat yet avoids potentially problematic correlations among the contextual predictors.

ANALYTIC STRATEGY AND PROCEDURES

The following analyses use hierarchical modeling procedures designed to account for the nested nature of multilevel sentencing data. Although the logic and necessity of hierarchical modeling in the study of criminal sentencing has been recently enumerated elsewhere (see Britt, 2000; Kautt, 2002; Ulmer and Johnson, 2004), a brief discussion of these procedures provides a useful analytical backdrop. Hierarchical linear modeling (HLM) techniques overcome several methodological and conceptual difficulties present in earlier analyses of contextual effects in sentencing. With individual criminal cases nested within judges, and judges nested within judicial districts, statistical dependency problems are likely to arise. Cases sentenced by the same judges and in the same districts are likely to have certain similarities. Statistically speaking, this means that residual errors are likely to be correlated within judges and within county-level courts, violating fundamental error assumptions of ordinary least squares regression and resulting in misestimated standard errors. Simply put, "misestimated standard errors occur with multilevel data when we fail to take into account the dependence among individual responses within the same organization. . . hierarchical linear models resolve this problem by incorporating into the statistical model a unique random effect for each organizational unit" (Raudenbush and Bryk, 2002: 100). Statistical procedures that do not account for the hierarchical nesting of data risk the artificial inflation of statistical power at higher levels of analysis by failing to properly adjust the degrees of freedom to the appropriate sample size. The degrees of freedom for higher levels of analysis are not a product of the total number of cases, but rather a function of the number of level two (judges) or level three (counties) units in the data.

In addition, HLM also allows one to model the heterogeneity of regression coefficients that can occur when relationships between individual characteristics and outcomes vary across aggregate units. For example, the effect of being a minority offender may differ across judges. Multilevel modeling techniques allow one to model this variation by allowing both slopes and intercepts to vary across macro-level units. Overall, then, HLM corrects standard errors by accounting for the nested nature of sentencing data, adjusts statistical significance tests to reflect the appropriate degrees of freedom, and provides the researcher with important tools for assessing the random variation in individual-level sentencing factors across judges and counties. A three-level hierarchy characterizes the present data, with individual criminal cases nested within judges and judges nested within county-level judicial districts. Models examining incarceration were estimated with hierarchical logistic regression and those examining sentence length used hierarchical linear models. All variables were centered on their grand means and results

reported are based on unit-specific models using robust standard errors (Raudenbush and Bryk, 2002: 276–80).

The analysis begins by estimating unconditional models. These models produce estimates of the relative amount of sentencing variation that occurs at the individual, judge, and county levels of analysis, providing useful insights into the relative importance of judge and county social contexts in criminal sentencing. Individual-level predictors are then grand-mean centered and added to the models to assess the degree to which judge- and county-level variations are accounted for by compositional differences. This stage of the analysis also provides important information on the extent to which individual-level factors vary significantly across judges and county-level courts. Next, full three-level hierarchical models are estimated to investigate the direct effects of specific judge- and county-level contextual influences on criminal sentencing outcomes.[10] Finally, cross-level interactions are specified to investigate theoretical predictions regarding interactions between individual offender- and judge-level characteristics in sentencing. Overall, each stage of the analysis provides additional information regarding the importance of judge and county contexts in the courtroom decision-making process.

FINDINGS

Table 1 presents the descriptive statistics. Not surprisingly, these results demonstrate that the vast majority of sentencing judges in Pennsylvania are male and white. About 85 percent of all judges are men, and only about 7 percent are minorities. Translated into concrete numbers, this means that of the 303 judges in the analyses, forty-six were female and twenty-two were minorities. Relative to existing research, these are substantial numbers. Table 1 also demonstrates that judges experience different caseload compositions and pressures. This suggests that they may also experience different organizational demands to effectively expedite criminal case processing. Moreover, the types of crime that different judges sentence vary widely, with some judges sentencing as many as 60 percent violent crimes, and others sentencing drug cases almost exclusively.

---

10. Whereas individual-level variables are allowed to vary across judges and counties, judge level effects are constrained to be fixed across counties. This was in line with the present focus on individual variations in sentencing across judges and courts. It was also necessary to ensure a sufficient number of counties for which a unique regression equation could be estimated. Still, future work investigating the potential ways judge characteristics vary across county court contexts could prove interesting.

Significant variations also exist across Pennsylvania's criminal courts. The majority of courts are located in small, rural counties, though an almost equal number of cases are sentenced in large, medium, and small court jurisdictions. Some courts appear to strictly adhere to guideline prescriptions, and others mete out departure sentences for nearly one in four cases. Similar variations are found in the availability of local resources, with some courts experiencing plentiful local jail space and others being virtually without it. The county-level social environments surrounding the courts also vary markedly, including notable differences in their socioeconomic status, political ideology, and the racial composition of their populations. To the extent that sentencing outcomes vary across judges and jurisdictions, then, they may be tied to variations in judicial backgrounds and caseloads as well as courtroom social contexts.

INDIVIDUAL SENTENCING VARIATIONS

Table 2 presents results from the three-level unconditional models of incarceration and sentence length. Results from these models suggest that approximately 5 percent of the total variation in the likelihood of incarceration can be attributed to differences between judges, and an additional 5 percent is accounted for by counties.[11] Similarly, the results for sentence length indicate that about 6 percent of the total variance is attributable to judges, compared to about 7 percent for counties. Overall, about equal amounts of variation appear to exist across sentencing judges and across county-level courts.

Table 3 presents results for the individual-level influences on incarceration and sentence length. Table 4 presents those for examining the random variation of these effects across judges and counties. The first set of results is consistent with existing research on individual effects in sentencing. As expected, legal variables, such as offense severity, prior criminality, and presumptive sentence recommendation, dominate these models. However, extralegal effects are also noteworthy, indicating potentially important individual disparities: female offenders are .61 times less likely to be incarcerated, black and Hispanic offenders are respectively 56 and 48 percent more likely to be incarcerated, and offenders convicted at trial are almost twice as likely to receive an incar-

---

11. Because the incarceration outcome is a dichotomous variable, it lacks a meaningful individual-level variance component. If the level 1 model is conceived of in terms of a latent variable (see Snijders and Bosker, 1999, cited in Raudenbush and Bryk, 2002: 334), though, then the level 1 random effect can be assumed to have a standard logistic distribution with a mean of 0 and variance = $\pi^2/3$. Under this assumption the intraclass correlation can be estimated, though its meaningfulness depends on the validity of the underlying distributional assumptions.

Table 1. Descriptive Statistics for Individual-, Judge-, and Court-Level Predictors

Dependent variables		Mean	SD	Min	Max
In-out incarceration	N=148,590	.55	.50	0.0	1.0
Ln sentence length	N=79,333	.78	1.99	-4.6	6.5
Individual-level variables					
Sentence year (1999)		.48	.50	0.0	1.0
Offense severity		3.46	2.17	1.0	14.0
Prior criminality		1.31	1.84	0.0	8.0
In-out presumptive sentence		.30	.46	0.0	1.0
Length presumptive sentence		3.68	10.18	0.0	240.0
Mandatory applied		.23	.42	0.0	1.0
Offender age		31.50	10.23	12.0	99.0
Male offender (reference)		.82	.38	0.0	1.0
Female offender		.17	.37	0.0	1.0
White offender (reference)		.62	.49	0.0	1.0
Black offender		.28	.45	0.0	1.0
Hispanic offender		.06	.24	0.0	1.0
Violent offense		.13	.33	0.0	1.0
Property offense		.21	.41	0.0	1.0
Drug offense		.20	.40	0.0	1.0
Other offense (reference)		.46	.50	0.0	1.0
Non-negotiated plea (reference)		.17	.38	0.0	1.0
Negotiated plea		.65	.48	0.0	1.0
Trial		.04	.20	0.0	1.0
Judge-level variables	N=303				
Judge age		55.97	6.58	42.0	75.0
Female judge		.15	.36	0.0	1.0
Minority judge		.07	.26	0.0	1.0
Married judge		.69	.46	0.0	1.0
Miltary experience		.25	.44	0.0	1.0
Judicial tenure		11.91	8.45	0.0	42.0
Prosecutorial experience		.35	.48	0.0	1.0
Judicial caseload pressure		2.45	2.17	0.1	13.3
Percentage violent cases		14.09	10.29	0.0	60.6
Percentage property cases		20.32	6.88	2.2	41.9
Percentage drug cases		18.28	11.39	0.0	96.4
Court-level variables	N=60				
Large court		.03	.18	0.0	1.0
Medium court (reference)		.23	.43	0.0	1.0
Small court		.73	.45	0.0	1.0
Available jailspace		18.41	12.10	2.6	68.3
Guideline departure rate		7.38	4.38	1.2	24.6
Trial rate		1.61	1.10	0.0	5.0
Percentage unemployed		5.36	1.46	2.6	8.2
Percentage Hispanic		1.54	1.75	0.3	7.3
Percentage Republican		44.46	8.12	16.0	59.8

ceration sentence. Similar though less pronounced disparities emerge for sentence length.

Table 2. Three-Level Unconditional Hierarchical Models of Incarceration and Sentence Length

Incarceration			Ln Sentence Length		
Fixed effects	b	SE	Fixed effects	b	SE
Intercept	.36	.07***	Intercept	.62	.07***
Random effects	Variance	SD	Random effects	Variance	SD
Level 1	—	—	Level 1	3.23	1.80***
Level 2	.17	.41***	Level 2	.24	.49***
Level 3	.20	.44***	Level 3	.24	.49***
Between-judge proportion of variance			Between-judge proportion of variance		
	.05			.06	
Between-county proportion of variance			Between-county proportion of variance		
	.05			.07	

Note: Intraclass correlations for incarceration are based on the assumption that the level 1 random effect has a variance = $\pi^2/3$ (see footnote 11).
† $p \leq .10$ * $p < .05$ ** $p < .01$ *** $p < .001$

More germane are the findings in table 4. The significant variance components for the level two and level three model intercepts indicate that the likelihood and length of incarceration continue to vary across judges and courts respectively, even after accounting for differences in individual-level sentencing factors. Combining the variance components in table 4 with the coefficients in table 3 provides important insights into the magnitude of these variations. For an average offender (that is, at the mean of all explanatory variables), a one standard deviation change in the level two intercept, representing variation in sentencing across two-thirds of judges, results in the overall probability of incarceration varying between .63 and .80. Similarly, this probability varies between .54 and .86 across two-thirds (that is, one standard deviation) of county courts. This indicates that the likelihood of incarceration for similarly situated offenders is not only a function of their particular biographical and case characteristics, but also the result of the particular judge and county court in which the sentencing decision occurs.

Moreover, the results in table 4 demonstrate that for both incarceration and sentence length, the effects of virtually all the explanatory factors vary significantly across judges and courts.[12] This suggests that different judges

---

12. Although these models indicated that the Hispanic coefficient for incarceration varied significantly across both judges and counties, it was necessary to fix this effect in subsequent models to maximize the number of aggregate units for which a unique regression equation could be estimated and to facilitate model convergence.

weigh the importance of individual offense and offender characteristics differently, and that the influence of these factors also varies significantly across county contexts. Of particular interest is the effect of the presumptive guideline recommendation on the odds of incarceration. For two-thirds of judges, this odds ratio varies between 1.02 and 2.51, and for two-thirds of courts, between .83 and 3.09. If Pennsylvania's guidelines create greater uniformity in sentencing, as they are purported to do, one would expect the recommended sentence to exert similar influence across judges and courts. These results, however, clearly suggest this is not the case—different judges and courts place varying emphasis on the importance of following the presumptive recommendation to incarcerate (Johnson, 2005).

It is also interesting to note the degree to which the mode of conviction effect varies across judges and courts. On average, a trial conviction is associated with an 83 percent increase in an offender's sentence length. For two-thirds of judges, however, this effect varies between a 62 and a 107 percent increase. The variation across courts is even greater, increasing the final sentence length between 1.46 and 2.29 times for two-thirds of jurisdictions. Whereas a trial conviction is consistently associated with longer sentence lengths, then, the magnitude of the trial penalty depends considerably on the sentencing judge and court of conviction. To varying degrees, the other individual-level factors exert similar random effects across judges and courts, suggesting that the relative emphasis and interpretation of basic focal concerns in sentencing fundamentally depend on characteristics of the sentencing judge as well as elements of the courtroom community.

EXPLAINING SENTENCING VARIATIONS

Table 5 presents the results for the influence of judge- and county-level factors in sentencing, after controlling for individual case and offender characteristics. Results from these analyses indicate that several contextual factors exert direct effects on the likelihood and length of incarceration. As hypothesized, minority judges are somewhat less punitive, being .71 times as likely to sentence convicted offenders to incarceration, and meting out sentences that are 6 percent shorter than white judges. Though consistent with some prior findings (Spohn, 1990a), these results are contrary to recent work examining judge effects in Pennsylvania (Steffensmeier and Britt, 2001). This may reflect both differences in analytical sophistication and that judge characteristics tend to be related to aggregate court characteristics not controlled for in other studies.

This is not surprising given the limited number of Hispanic offenders in these data as well as their geographical concentration in relatively few counties.

Table 3. Three-Level HLM Random Coefficient Models—Individual-Level Fixed Effects

Individual-level factors	Incarceration			Ln Sentence Length	
	b	SE	Odds	b	SE
Constant	.98	.12	—***	.90	.04**
Guideline factors					
Year of sentence	.02	.03	1.02	.00	.02
Offense severity	.51	.02	1.66***	.65	.02***
Prior criminality	.46	.03	1.59***	.42	.02***
Presumptive sentence	.47	.09	1.60***	-.02	.00***
Mandatory applied	3.44	.33	31.24***	.04	.15
Offender factors					
Offender age	-.02	.00	.99***	.00	.00
Offender gender					
Male offender (reference)					
Female offender	-.49	.04	.61***	-.37	.03***
Offender race					
White offender (reference)					
Black offender	.44	.05	1.56***	.03	.01†
Hispanic offender	.39	.05	1.48***	.07	.03*
Other race-ethnicity	.05	.10	1.05	-.08	.06
Offense factors					
Offense type					
Violent offense	.34	.08	1.41***	-.07	.04†
Property offense	.22	.07	1.25***	.07	.03*
Drug offense	.10	.10	1.11	.02	.04
Other offense (reference)					
Case-processing factors					
Mode of conviction					
Non-negotiated plea (reference)					
Negotiated plea	-.15	.06	.86***	-.07	.03*
Trial	.67	.11	1.95***	.60	.04***
Selection bias correction factor				2.15	.26***
N		148,590		79,333	

† $p \leq .10$ * $p < .05$ ** $p < .01$ *** $p < .001$

Whereas the effect of judge's race was noteworthy, the influence of judge's gender was minimal. The age of the judge was significant, however, with older judges being less likely to incarcerate convicted offenders and sentencing them to shorter periods of confinement. Interestingly, prior military experience exerted a positive influence on the odds of incarceration, whereas the tenure of the judge was only marginally associated with increased sentence severity for both outcomes. Theoretically, military experience might reflect underlying judicial attitudes that are more punitive and crime-control oriented (Kritzer,

1978), and increased tenure on the bench might harden judges to the plight of the offender (Hogarth, 1971).

Table 4. Three-Level HLM Random Coefficient Models—Individual-Level Random Effects

	Incarceration			Ln Sentence Length		
Judge-level random effects	Variance	df	χ^2	Variance	df	χ^2
Level 1 intercept				1.209	—	—
Level 2 intercept	.182	171	555.0***	.024	184	505.7***
Offense severity	.015	171	382.8***	.003	184	465.0***
Prior criminality	.015	171	865.8***	.002	184	429.5***
Presumptive sentence	.204	171	364.7***	.000	184	360.8***
Violent offense	.155	171	433.2***	.015	184	248.1***
Property offense	.179	171	545.7***	.034	184	300.0***
Drug offense	.159	171	604.3***	.033	184	315.0***
Black	.044	171	254.1***	.009	241	298.8**
Hispanic	.079	171	231.8***	—	—	—
Female	.068	171	322.2***	.020	184	293.0***
Age	.000	171	283.9***	.000	184	312.7***
Negotiated plea	.199	171	579.1***	.022	184	324.4***
Trial	.267	171	262.8***	.015	184	244.3**
County-level random effects						
Level 3 intercept	.684	57	902.0***	.057	55	554.131***
Offense severity	.025	57	271.14***	.005	55	303.01***
Prior record	.032	57	381.33***	.003	55	261.75***
Presumptive sentence	.432	57	243.07***	.000	55	156.11***
Violent offense	.314	57	277.94***	.084	55	337.69***
Property offense	.304	57	269.64***	.067	55	275.50***
Drug offense	.492	57	377.42***	.091	55	289.51***
Black	.075	57	203.73***	—	—	—
Hispanic	.277	57	99.13***	.019	55	82.55**
Female	.037	57	97.98***	.023	55	137.57***
Age	.000	57	217.43***	.000	55	149.06***
Negotiated plea	.057	57	93.36**	.025	55	127.68***
Trial	.241	57	90.55**	.050	55	134.44***

† $p \leq .10$ * $p < .05$ ** $p < .01$ *** $p < .001$

Although the overall caseload pressure of the sentencing judge exerted insignificant effects on both the likelihood and the length of incarceration, the caseload composition of the judge demonstrated small but significant effects on sentencing decisions. Judges who sentenced greater proportions of property crimes were marginally more likely to incarcerate offenders, and heavy violent and drug caseloads led to significantly longer sentences. On the surface, these effects appear quite small, but their cumulative impact is potentially consequential. A one standard deviation increase in a

judge's violent caseload, for instance, results in about a 10 percent increase in sentence length. The present results suggest potentially interesting findings for judicial caseload factors, but additional research is necessary to further explore these relatively unexamined influences.

Table 5. Full Three-Level HLM Random Coefficient Models

	Incarceration			Ln Sentence Length	
	b	SE	Odds	b	SE
Intercept	1.00	.20	—***	.94	.03
Judge-level factors					
Judge age	-.02	.00	.98***	-.003	.00*
Female judge	-.06	.13	.94	-.02	.02
Minority judge	-.33	.06	.72***	-.06	.03*
Married judge	.04	.06	1.04	-.01	.02
Judicial military experience	.12	.06	1.13*	.03	.02
Judge tenure on bench	.01	.00	1.01†	.002	.00†
Prosecutorial experience	-.07	.05	.94	.02	.01
Judicial caseload pressure	.02	.02	1.02	-.01	.00
Judge violent caseload	.01	.01	1.01	.01	.00***
Judge property caseload	.01	.01	1.01†	.00	.00
Judge drug caseload	.01	.01	1.01	.005	.00***
County-level factors					
Large court	-.73	1.04	.48	-.15	.08†
Small court	.60	.21	1.83*	.07	.05
Available jail space	.02	.01	1.02*	.00	.00
Guideline departure rate	-.05	.02	.95*	.00	.01
Trial rate	-.02	.07	.98	.02	.02
Percentage unemployed	.01	.07	1.01	.02	.02
Percentage Hispanic	.01	.07	1.01	.02	.01*
Percentage Republican	.00	.02	1.00	.00	.00
Individual R-squared		—			.63
Judge R-squared		.05			.24
County R-squared		.51			.14
N		148,590		79,333	

Note: These models also include all individual-level control variables presented in table 3. They are not reported here because individual-level findings were little changed when judge- and county-level factors were included.
† $p \leq .10$ * $p < .05$ ** $p < .01$ *** $p < .001$

Select hypotheses also predicted significant influences from county-level courtroom factors, after accounting for judge-level covariates. In line with courtroom community theory, a pattern of sentencing leniency emerged in association with large courts, whereas small courts were generally more punitive. In particular, offenders sentenced in small courts were 83 percent more likely to be incarcerated than those sentenced in

midsized courts, with the difference between small and large courts being even greater. Sentence lengths in larger courts were also about 15 percent shorter than in medium courts (though marginally significant), and they were 22 percent shorter than in small courts. Overall, court size demonstrated important effects on sentencing severity, though interestingly, its influence on sentence length was reduced substantially and became insignificant in additional models without the correction for selection bias.[13]

Both the guidelines departure rate of the court and local jail capacity also demonstrated significant influences on the likelihood of incarceration. Courts with higher departure rates were less likely to incarcerate offenders. High departure rates may reflect courtroom norms encouraging greater reliance on substantive sentencing considerations. Given evidence of judicial perceptions of undue guideline severity for serious offenses (Kramer and Ulmer, 2002), it is not surprising that this led to reduced use of incarceration. As the focal concerns perspective suggests, courtroom environments characterized by greater local jail capacity were also more likely to incarcerate otherwise comparable offenders. Finally, consistent with racial threat perspectives (for example, Blalock, 1967; Liska, 1992), communities with larger Hispanic populations were characterized by slightly longer sentences. Overall, these findings substantiate theoretical expectations. They are noteworthy in that they replicate and extend recent research on courtroom contexts by simultaneously incorporating various theoretically important judge-level predictors of sentencing.

Finally, in addition to the direct impact of judge and courtroom social contexts, this study also hypothesized select cross-level interactions between judge and offender characteristics. Table 6 presents the results from these analyses. Although in the expected direction for incarceration, the interaction between offender and judge gender produced small and

---

13. It is also worth commenting on the sizeable but statistically insignificant effect for large courts on incarceration. Overall, few differences emerged for models comparing normal and robust standard errors. The effect of large courts, however, represented an important exception. Consistent with prior work (for example, Ulmer and Johnson, 2004), this effect was strong and statistically significant when normal standard errors were applied ($b = -.73$; SE = .31), but became statistically insignificant with robust standard errors. This may be due to the fact that only two counties (Philadelphia and Allegheny) qualified as large courts. Supplemental analysis demonstrated that when the three largest medium-sized counties were coded as large courts, the robust standard error was notably reduced. However, the coefficient for large courts was also substantially reduced supporting the initial decision to code these counties as medium-sized courts. Further examination of a continuous measure of court size (based on number of judges in the county) bolstered arguments for its overall importance, indicating that counties with more judges were significantly less likely to incarcerate offenders ($b = -0.05$; SE = 0.01).

statistically insignificant effects. Overall, there was little evidence that the gender of the judge had any significant direct or indirect influence on judicial sentencing behavior.[14] However, strong evidence emerged for the racial interactions between judges and offenders. Minority judges were significantly less likely to incarcerate both black and Hispanic offenders. Overall, minority offenders were about 50 percent more likely to be incarcerated; however, this disparity was reduced by more than half when a minority judge presided. Surprisingly, though, minority judges also sentenced black offenders to slightly longer terms of incarceration. These findings suggest the potential importance of the judge's racial background, but also highlight the complex and at times counterintuitive nature of contextual interactions in sentencing.

Table 6. HLM Cross-Level Judge Interaction Models

	Incarceration			*Ln Sentence Length*	
Gender Interactions	b	SE	Odds	b	SE
Intercept	1.00	.11	—***	.94	.03***
Female judge	-.06	.08	.94	-.01	.02
Female	-.48	.04	.62***	-.37	.03***
Female*female judge	.08	.07	1.08	-.02	.03
Race-ethnicity interactions					
Intercept	1.00	.20	—***	.94	.03***
Minority judge	-.28	.06	.75***	-.09	.02***
Black	.42	.05	1.53***	.03	.01†
Black*minority judge	-.27	.11	.76*	.07	.02**
Hispanic	.42	.07	1.51***	.07	.03*
Hispanic*minority judge	-.29	.09	.75***	.06	.05
Caseload composition interactions					
Intercept	1.00	.16	—***	.94	.03***
%Violent	.01	.01	1.01	.01	.00***
Violent	.32	.11	1.38**	-.08	.04*
Violent*%violent	-.02	.00	.98***	-.01	.00***
N		148,590		79,333	

Note: Cross-level interaction models include all individual- and contextual-level control variables.
† $p \leq .10$ * $p < .05$ ** $p < .01$ *** $p < .001$

Theoretical predictions regarding judicial caseload factors suggested that individual sentencing decisions may be influenced by the type of

---

14. Additional examinations of other judicial gender interactions produced similar null findings. For instance, contrary to some prior research (Steffensmeier and Hebert, 1999), there was no evidence that female judges sentenced minority offenders any differently from male judges. Interestingly, though, supplemental analyses did suggest that minority judges sentenced female offenders to significantly shorter sentence lengths.

caseload experienced by the judge. Convincing evidence of this emerged, with the effect of a violent conviction being significantly conditioned by the violent caseload of the sentencing judge for both outcomes. Judges who sentenced a higher proportion of violent crimes were less likely to incarcerate violent offenders, and, when they did, sentenced them to shorter periods of confinement. For example, a one standard deviation increase in a judge's violent caseload reduced the odds ratio for violent crime on incarceration from 1.38 to 1.15. This suggests that judicial caseloads have significant conditioning effects on individual-level judicial considerations in sentencing.

Examining the collective impact of judge- and courtoom-level predictors is instructive, given their notable influences on criminal sentencing. For example, the combined influence of age, race, and gender is a significant factor in the likelihood of incarceration. Holding all other explanatory variables at their means, the likelihood of incarceration for cases sentenced by a younger (that is, one standard deviation below the mean) white male judge is .76, whereas for an older (one standard deviation above the mean) minority female judge it is .63. These differences become even more pronounced when courtroom contexts are also considered. The overall probability of incarceration for an offender sentenced by the younger white male judge in a small county is .79, and for the older minority female judge in a large court a dramatically lower .35. These results are particularly interesting given that minority and female judges in Pennsylvania tend to congregate in the relatively few large, urban courts. Similar variations were observed for predicted sentence lengths, but the substantive magnitude of these effects was not impressive. For instance, the expected logged sentence length for a young white male judge was 1.00 (exp(1.00) = 2.72 months), whereas for an older minority female judge it was .81 (exp(.81) = 2.25 months). Supplementary models using an unlogged measure of sentence length produced similarly small differences.

Of course the aforementioned predictions are based on artificial constraints that assume the same "average" offender is being processed while select judge- and county-level factors are manipulated. In reality, judge and county court contexts are likely tied to one another in more intricate and complex ways that are not captured in these simple estimates. Although additional, thoughtful research is needed to further explore this important nexus, the above estimates provide a useful though preliminary illustration of the potentially important and complementary influences that judge- and county-level factors exert on individual sentencing decisions across courtroom decision-making contexts.

DISCUSSION

This study examined several theoretical hypotheses regarding the influence of judge- and county-level contexts in criminal sentencing. Collective support for these expectations is summarized in table 7, with statistical significance levels provided for predicted effects. Overall, the findings reinforce the notion that judges are primarily influenced by individual sentencing factors; however, results from this study also suggest nontrivial variations occur across judges and county courts, even after accounting for differences in individual cases. Moreover, the effects of individual sentencing factors also varied across contexts, at times resulting in sizeable sentencing differences.

Table 7. Support for Theoretical Predictions of Contextual Effects in Sentencing

		Supported	
Description of hypotheses		*In-Out*	*Length*
1.	Sentence severity varies across both judges and county courts.	Yes***	Yes***
2.	Individual level sentencing effects vary across judges and courts.	Yes***	Yes***
3.	Minority judges sentence more leniently than white judges.	Yes***	Yes*
4.	Female judges sentence more leniently than male judges.	No	No
5.	Judges with heavier caseloads sentence more leniently.	No	No
6.	Judges with heavier violent caseloads sentence more leniently .	No	No
7.	Small courts sentence more severely than large courts.	Yes**	Yes
8.	Higher departure rates result in more lenient sentencing.	Yes*	No
9.	Available jail capacity increases the likelihood of incarceration.	Yes*	—
10a.	Minority judges sentence black offenders more leniently.	Yes*	No
10b.	Minority judges sentence Hispanic offenders more leniently.	Yes***	No
11.	Male judges sentence female offenders more leniently.	Yes	No
12.	Violent crime caseload conditions the violent crime effect.	Yes***	Yes***

† $p \leq .10$ * $p < .05$ ** $p < .01$ *** $p < .001$

Attempts to explain these contextual variations suggested that minority judges were significantly less likely to incarcerate offenders, as were older judges, and that the combined influences of judge and court contexts were especially great, particularly for the incarceration decision. In line with courtroom community theory, offenders sentenced in small courts were more likely to be incarcerated, whereas offenders in large courts received marginally longer sentences. Offenders sentenced in counties with more available jail space were also more likely to be incarcerated. As the focal

concerns perspective argues, judges and other courtroom actors are keenly aware of and influenced by local resources such as jail capacity. Finally, courts with higher overall departure rates were less likely to incarcerate otherwise comparable offenders. This may indicate that courtroom environments that emphasize greater substantive rationality are less punitive. Although these results for courtroom contextual effects are in line with existing research, they extend it significantly by substantiating these effects beyond the influence of judicial characteristics in sentencing.

Perhaps most important were the findings for cross-level interactions between judge and offender characteristics. No research to date has used multilevel analytical techniques to investigate these relationships. Although the gender of the judge had little impact on sentencing, the race of the judge conditioned the influence of offender's race or ethnicity. Minority judges were significantly less likely than white judges to incarcerate black and Hispanic offenders, but still incarcerated them more often than they did white offenders. This may lend credence to recent arguments that an increased minority presence on the bench will attenuate racial and ethnic disparities (Walker and Barrow, 1985). However, minority judges also sentenced black offenders to longer terms, suggesting the relationship between judge and offender racial backgrounds is complicated and merits additional investigation.

Although the caseload pressure of the judge was unrelated to sentencing outcomes, it is important to note two caveats. First, this measure was limited to criminal cases, thus did not capture potentially important differences in judges' civil caseloads. Second, given the predominance of cases convicted through guilty pleas, the caseload of the prosecutor may be the more germane measure of courtroom efficiency. Because the ratio of prosecutors to judges varies across courts, judicial caseloads may not accurately reflect prosecutorial caseloads. Results for the caseload composition of the judge proved more fruitful. Judges faced with heavier violent caseloads were less punitive toward violent offenders. As Emerson (1983) argued, individual sentences are determined in the context of the overall flow of cases. Judges faced with frequent violent offenders may therefore be desensitized to the relative severity of violent crimes, resulting in relative leniency at sentencing. Future studies should further explore the importance of judicial caseload factors, examining their influence for additional types of crime in addition to measures of overall caseload severity across sentencing judges.

CONCLUSION

This study is an attempt to further current research on contextual effects in criminal sentencing by integrating analyses of judge

characteristics and courtroom contexts. In pursuit of this goal, the study examined a broad range of judicial background factors and county-level social contexts, accounting for a host of relevant individual-level influences. Results from this investigation support theoretical arguments for the complexity of courtroom decision making and highlight the role that social contexts play in this multifaceted process.

Overall, these findings demonstrate that the criminal sentencing decision-making process is jointly influenced by individual case characteristics, judicial background factors, and county-level contextual influences. Moreover, the study demonstrates that even under presumptive guidelines, sentencing outcomes continue to vary significantly across courtroom actors and their courts. These variations, then, raise important questions regarding contextual disparities in sentencing under statewide, presumptive guidelines. As some scholars have noted, the sentencing guidelines are an attempt to restore formal rationality to the criminal process. Because these recommendations are filtered through individual courtroom actor interpretations, however, and because they are colored by informal, locally varying courtroom norms, it is not surprising that they have failed to eliminate judge and court variation in sentencing (Savelsberg, 1992; Ulmer and Kramer, 1996).

These variations are reflected not only in differences in overall sentencing severity, but also in the relative weight attached to virtually all individual-level sentencing considerations. This suggests that the relative emphasis and interpretation of focal concerns in sentencing depend on both judicial and courtroom community contexts. That the effect of the presumptive guidelines recommendation varied significantly across judges and courts, for instance, highlights the importance of local courtroom actor interpretations of formal sentencing recommendations. Similarly, that the influence of individual offender characteristics, such as race-ethnicity, varied across judges and counties suggests that extralegal sentencing disparities may be judge and county specific, but not any less real under sentencing guideline systems. Future research attempting to evaluate the effectiveness of sentencing guidelines therefore needs to do so with a sensitivity to judge and context-specific disparities.

Although this study extends the isolated research on judge and county contexts in important ways, it is not without its limitations. Information on theoretically important variables, such as offender socioeconomic status and victim characteristics, as well as judicial political affiliation and social class, was not available. Still, this study incorporated a wider range of judge and county factors than most prior studies. Despite this, considerable variation between judges and between counties remained unexplained in both incarceration and sentence length. Collectively, the judicial background factors examined accounted for only 5 percent and 24

percent of the judge variation in incarceration and sentence length respectively (see table 5). While the courtroom context indicators accounted for 51 percent of the between county variation in incarceration, they only accounted for 14 percent of the variation in sentence length. The task of future research on contextual effects in sentencing is to search for additional theoretical measures that better tap the sources of disparity between judges and between courts. Judicial caseload factors and courtroom departure rates are a step in this direction. Additional factors such as career paths to the judgeship may also prove fruitful. Direct measures of judges' punitive philosophies and attitudes towards crime-control would almost certainly improve one's ability to explain judicial disparities in sentencing (Hogarth, 1971). More direct measures of courtroom community norms and workgroup expectations are also needed to better tap into the elusive concept of local courtroom culture (Ulmer and Johnson, 2004).

In addition to elaborating the theoretical scope, future research needs to begin to incorporate information on other courtroom actors in the sentencing process. In particular, a better sense of prosecutorial factors may improve our understanding of courtroom decision making considerably. If prosecutorial discretion is particularly influential in plea negotiation outcomes, as recent scholarship suggests (Johnson, 2003), then so-called judicial variations in sentencing may be due in part to unexamined variations in prosecutorial factors. Future research is thus needed that better specifies the influences of both additional courtroom actors and additional courtroom contexts.

Finally, because judicial background factors are inextricably linked to courtroom community characteristics, it is imperative to continue to investigate how multiple levels of social context jointly condition individual sentencing decisions. This is not to suggest that individual-level analyses of criminal sentencing are inconsequential. On the contrary, individual-level variables consistently have the strongest effects in sentencing and are robust even when contextual predictors are included. This study serves instead to highlight the subtle but important nature of contextual influences in sentencing across multiple levels of analysis. By continuing to tackle the complex interplay among individual-, judge-, and county-level correlates of sentencing severity, subsequent work will better inform our understanding of the subtleties inherent in the courtroom decision-making process.

REFERENCES

Albonetti, Celesta. 1991. An integration of theories to explain judicial discretion. *Social Problems* 38:247–66.

Albonetti, Celesta. 1997. Sentencing under the federal sentencing guidelines: An analysis of the effects of defendant characteristics, guilty pleas, and departures on sentencing outcomes for drug offenses. *Law and Society Review* 31:601–34.

Berk, Richard. 1983. An introduction to sample selection bias in sociological data. *American Sociological Review* 48:386–98.

Blalock, Herbert. 1967. *Toward a Theory of Minority Group Relations.* New York: Carpricorn Books.

Bontrager, Stephanie, William Bales, and Ted Chiricos. 2005. Race, ethnicity, threat and the labelling of convicted felons. *Criminology* 43(3): 589–622.

Britt, Chester. 2000. Social context and racial disparities in punishment decisions. *Justice Quarterly* 17(4): 707–32.

Bushway, Shawn, and Anne Piehl. 2001. Judging judicial discretion: Legal factors and racial discrimination in sentencing. *Law and Society Review* 35(4): 733–67.

Crawford, Charles, Ted Chiricos, and Gary Kleck. 1998. Race, racial threat, and sentencing of habitual offenders. *Criminology* 36(3): 481–512.

D'Alessio, Stewart, and Lisa Stolzenberg. 1997. The effect of available capacity on jail incarceration: An empirical test of Parkinson's Law. *Journal of Criminal Justice* 25(40): 279–88.

Dixon, Jo. 1995. The organizational context of criminal sentencing. *American Journal of Sociology* 100:1157–198.

Eisenstein, James, Roy Flemming, and Peter Nardulli. 1988. *The Contours of Justice: Communities and Their Courts.* Boston: Little, Brown.

Emerson, Robert M. 1983. Holistic effects in social control decision-making. *Law & Society Review* 17(3): 425–55.

Frazier, Charles E., and Wilbur Bock. 1982. Effects of court officials on sentence severity: Do judges make a difference? *Criminology* 20:257–72.

Gibson, James, L. 1983. From simplicity to complexity: The development of theory in the study of judicial behavior. *Political Behavior* 5:7–50.

Goldman, Sheldon. 1979. Should there be affirmative action for the judiciary? *Judicature* 62:488–94.

Granovetter, Mark. 1985. Economic action and social structure: The problem of embeddedness. *American Journal of Sociology* 91(3): 481–510.

Gruhl, John, Cassia Spohn, and Susan Welch. 1981. Women as policymakers: The case of trial judges. *American Journal of Political Science* 25:308–22.

Hawkins, Darnell. 1981. Causal attribution and punishment for crime. *Deviant Behavior* 2:207–30.

Heckman, James. 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables. *Annals of Economic and Social Measurement* 5:475–92.

Helms, Ronald, and David Jacobs. 2002. The political context of sentencing: An analysis of community and individual determinants. *Social Forces* 81(2): 577–604.

Hogarth, John. 1971. *Sentencing as a Human Process.* Toronto: University of Toronto Press in Association with the Centre of Criminology.

Holleran, David, and Cassia Spohn. 2004. On the use of the total incarceration variable in sentencing research. *Criminology* 42(1): 211–41.

Holmes, Malcolm D., Harmon M. Hosch, and Howard C. Daudistel. 1993. Judge's ethnicity and minority sentencing: Evidence concerning Hispanics. *Social Science Quarterly* 74(3): 496–506.

Johnson, Brian D. 2003. Racial and ethnic disparities in sentencing departures across modes of conviction. *Criminology* 41(2): 449–90.

Johnson, Brian D. 2005. Contextual disparities in guidelines departures: Courtroom social contexts, guidelines compliance, and extralegal disparities in criminal sentencing. *Criminology* 43(3): 761–96.

Kautt, Paula, M. 2002. Location, location, location: Interdistrict and intercircuit variation in sentencing outcomes for federal drug-trafficking offenses. *Justice Quarterly* 19(4): 633–71.

Kramer, John H., and Jeffery T. Ulmer. 2002. Downward departures for serious violent offenders: Local court corrections to Pennsylvania sentencing guidelines. *Criminology* 40(4): 897–933.

Kritzer, Herbert M. 1978. Political correlates of behavior of judges: A best case analysis. *Journal of Politics* 40:25–58.

Kritzer, Herbert M., and Thomas M. Uhlman. 1977. Sisterhood in the courtroom: Sex of judge and defendant in criminal case disposition. *Social Science Journal* 14:77–88.

Kulik, Carol, Elissa Perry, and Molly Pepper. 2003. Here comes the judge: The influence of judge personal characteristics on federal sexual harassment case outcomes. *Law and Human Behavior* 27(1): 69–86.

Kurlychek, Megan, and Brian D. Johnson. 2004. The juvenile penalty: A comparison of juvenile and young adult sentencing outcomes in criminal court. *Criminology* 42(2): 485–517.

Levin, Martin A. 1977. *Urban Politics and the Criminal Courts.* Chicago: University of Chicago Press.

Liska, Allen. 1992. *Social Threat and Social Control.* Albany: State University of New York Press.

Mustard, David B. 2001. Racial, ethnic, and gender disparities in sentencing: Evidence from U.S. federal courts. *Journal of Law and Economics* 44(1): 285–314.

Myers, Martha, and Susette Talarico. 1987. *The Social Contexts of Criminal Sentencing.* New York: Springer-Verlag.

Peterson, Ruth D., and John Hagan. 1984. Changing conceptions of race: Toward an account of anomalous findings of sentencing research. *American Sociological Review* 49:56–70.

Raudenbush, Stephen, and Anthony Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park, CA: Sage Publications.

Savelsberg, Joachim. 1992. Law that does not fit society: Sentencing guidelines as a neoclassical reaction to the dilemmas of substantivized law. *American Journal of Sociology* 97:1346–81.

Sellin, Thorsten. 1928. The Negro criminal: A statistical note. *Annals of the American Academy of Political and Social Science* 140:52–64.

Snijders, Tom, and Roel Bosker. 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling.* Thousand Oaks, CA: Sage Publications.

Songer, Donald R., and Kelley Crews-Meyer. 2000. Does judge's gender matter? Decision making in U.S. state supreme courts. *Social Science Quarterly* 81(3): 750–62.

Spohn, Cassia. 1990a. The sentencing decision of black and white judges: Some expected and unexpected similarities. *Law and Society Review* 24:1197–216.

Spohn, Cassia. 1990b. Decision making in sexual assault cases: Do black and female judges make a difference? *Women and Criminal Justice* 2(1): 83–105.

Steffensmeier, Darrell, and Chester Britt. 2001. Judge's race and judicial decision making: Do black judges sentence differently? *Social Science Quarterly* 82(4): 749–65.

Steffensmeier, Darrell, and Stephen DeMuth. 2000. Ethnicity and sentencing outcomes in U.S. federal courts: Who is punished more harshly? *American Sociological Review* 65:705–29.

Steffensmeier, Darrell, and Chris Hebert. 1999. Women and men policymakers: Does the judge's gender affect the sentencing of criminal defendants? *Social Forces* 77(3): 1163–196.

Steffensmeier, Darrell, Jeffery T. Ulmer, and John Kramer. 1998. The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black, and male. *Criminology* 36:763–98.

Stolzenberg, Ross M., and Daniel A. Relles. 1990. Theory testing in a world of constrained research design: The significance of Heckman's censored sampling bias correction for nonexperimental research. *Sociological Methods and Research* 18:395–415.

Sudnow, David. 1965. Normal crimes: Sociological features of the penal code. *Social Problems* 12(4): 255–64.

Uhlman, Thomas. 1978. Black elite decision making: The case of trial judges. *American Journal of Political Science* 22:884–95.

Ulmer, Jeffery T. 1997. *Social Worlds of Sentencing: Court Communities Under Sentencing Guidelines.* Albany: State University of New York Press.

Ulmer, Jeffery T., and Brian D. Johnson. 2004. Sentencing in context: A multilevel analysis. *Criminology* 42(1): 137–77.

Ulmer, Jeffery T., and John Kramer. 1996. Court communities under sentencing guidelines: Dilemmas of formal rationality and sentencing disparity. *Criminology* 35:306–32.

Walker, Thomas, and Deborah Barrow. 1985. The diversification of the federal bench: Policy and process ramifications. *The Journal of Politics* 47:596–617.

Welch, Susan, Michael Combs, and John Gruhl. 1988. Do black judges make a difference? *American Journal of Political Science* 32(1): 126–36.

Wheeler, Stanton, David Weisburd, and Nancy Bode. 1982. Sentencing the white collar offender: Rhetoric and reality. *American Sociological Review* 47(5): 641–59.

Zatz, Marjorie. 2000. The convergence of race, ethnicity, gender, and class on court decision making: Looking toward the 21st century. *Criminal Justice: The National Institute of Justice Journal* 3:503–52.

Brian D. Johnson is an assistant professor of Criminology and Criminal Justice at the University of Maryland. His research interests include contextual variations in criminal sentencing, racial and ethnic disparities in the justice system, and the use of advanced statistical modeling techniques to study criminal processes.

Appendix.   Statistical Coding and Descriptions of Variables

Dependent Variables	*Coding*	*Description*
In-out incarceration	1=Incarcerated	Dummy variable identifying cases that were sentenced to incarceration
Ln sentence length	Log of number of months	Natural logarithm of minimum number of months of incarceration

Independent Variables

Individual-level predictors

Year of sentence	1=1999	Dummy variable identifying cases sentenced in 1999 instead of 2000
Offense severity (OGS)	1-14	PCS Offense Gravity Score measured by 1-14 point scale
Criminal history (PRS)	0-8	PCS Prior Record Score measured by 0-8 point scale
Mandatory	1=Mandatory applied	PCS dummy variable identifying cases that received a mandatory minimum
In-out presumptive	1=Incarceration prescribed	Dummy variable identifying whether guidelines precribe incarceration
Length presumptive	Number of months	Continuous variable measuring minimum months of incarceration guidelines prescribe
Offense type	1=Violent offense 1=Property offense 1=Drug offense 0=Other offense (reference)	Dummy variables identifying cases convicted for violent crimes, property crimes and drug crimes, with other crimes serving as reference
Offender age	Number of years	PCS continuous variable measuring the age of the offender at the date of sentencing
Offender race	0=White (reference) 1=Black 1=Hispanic 1=Other race	PCS dummy variables identifying an offender's race-ethnicity as white, black, Hispanic, or other race, with white offenders serving as reference
Offender gender	1=Female	PCS dummy variable identifying the gender of the offender, with male serving as reference
Mode of conviction	0=Non-negotiated plea 1=Negotiated plea 1=Bench trial 1=Jury trial	PCS dummy variables identifying cases convicted through negotiated pleas, bench trials, and jury trials, with non-negotiated pleas serving as reference
Selection bias factor	Hazard rate	Control for selection bias calculated following Heckman (1976) and Berk (1983)

Appendix. Statistical Coding and Descriptions of Variables (continued)

Independent Variables	Coding	Description
Judge-level predictors		
Judge age	Number of years	Continuous variable measured as age of judge in 2000
Judge gender	1=Female	Dummy variable identifying cases sentenced by a female judge
Judge race	1=Minority	Dummy variable identifying cases sentenced by a minority judge
Marital status	1=Married	Dummy variable identifying cases sentenced by a married judge
Military experience	1=Military	Dummy variable identifying cases sentenced by judges with experience in military
Judicial tenure	Number of years	Continuous variable measured as years serving on Pennsylvania Court of Common Pleas
Prosecutorial experience	1=Prosecutor	Dummy variable identifying cases sentenced by judges with prosecutorial legal expereince
Judicial caseload pressure	Number of cases/2	Average number of criminal cases sentenced by a judge in a year (subsequently divided by 100)
Judicial case composition	% violent % property % drug	Three continuous variables measuring percentage of total criminal cases sentenced by each individual judge for violent crimes, property crimes, and drug crimes
County-level predictors		
Court size	1=Large court 0=Medium court (reference) 1=Small court	Three dummy variables identifying large, medium, and small courts, based on number of cases and judges in each county, with medium courts serving as reference
Available jail space	Number of jail beds	Continuous measure of jail beds divided by number of cases sentenced in county
Guideline departure rate	% of cases	Continuous measure of percentage of cases in county sentenced outside sentencing guidelines
Trial rate	% of cases	Continuous measure of percentage of cases convicted through jury trials in the county
Socioeconomic climate	% unemployed	Continuous measure of percentage of county population that is unemployed
Racial composition	% Hispanic	Continuous measure of percentage of county population identified as Hispanic
Political climate	% Republican	Continuous measure of percentage of county population identified as Republican