Unwarranted disparity in sentencing continues to be a problem in criminal courts. Sentencing, a final step in the multi-stage process of criminal justice, is seen as such an important component of ensuring consistency that legislative policy has created guidelines to control judicial decision-making. Nonetheless, research shows that disparity under sentencing guidelines persists due to influences external to the law—referred to as substantive rationality. Sentencing research overwhelmingly utilizes quantitative analysis of official court data to produce its conclusions about the influences of disparity. The current study builds upon past research by using qualitative analysis of interviews with 20 judges in Michigan. The recent shift from mandatory to advisory sentencing guidelines and the scope of disparity in Michigan make it a substantively rich site for investigation. This analysis found several themes in judges’ perspectives on causes of disparity, including: (1) localization, organizational resources, and court policy; (2) individual judicial philosophy and decision-making; (3) influence from pre-sentencing reports and appellate courts.

Keywords: sentencing; disparity; judges; qualitative research methods; court communities

INTRODUCTION

The goal of uniform sentencing within sentencing guidelines has, in many ways, proven to be “unfit for society” (Savelsberg, 1992). Across numerous studies, we see how the legal attempt to control disparity by restricting judicial decision-making faces durable resistance from individualized, organizational, political, and cultural forces (Spohn, 2009). Importantly, unwarranted sentencing disparities are frequently cited as an elemental source of the social inequalities that persist in the US carceral system (Alexander, 2010). Over the last several decades, theoretical and empirical research on sentencing disparity shows that its causes are multi-level and develop across multiple stages in the criminal justice process (Spohn, 2009; Ulmer, 2012; Kutateladze et al., 2014). The purpose of this study is to examine how windows of disparity develop in a guidelines system with goals of formal rationality (Savelsberg, 1992; Ulmer & Kramer, 1996).

Although the problem of disparities is a complex one, at its core, it is simple: guidelines limit discretion, but they do not result in uniform outcomes. The “courts as communities” perspective, developed by Eisenstein, Flemming, and Nardulli (1988), provides one theoretical

1 University of North Carolina-Wilmington, Department of Sociology and Criminology, 125 Bear Hall Wilmington NC 28403, smithjm@uncw.edu
approach toward understanding disparities and posits that different sentencing practices result from varying legal, cultural norms, politics, and organizational arrangements. This perspective and other complementary theories on sentencing, contextualize the findings, which both support and refine past research.

Michigan serves as the site of this study for theoretical and policy-based reasons. First, the Council of State Governments (2014: 7) recently illustrated wide ranges of disparity in Michigan, stating that sentences are “all over the map” and “people with similar criminal histories who are convicted of similar crime receive significantly different sentences.” Additionally, of its felony dispositions, Michigan incarcerates at a higher percentage compared to other guidelines states, especially through jail. However, this varies widely by jurisdiction. As Figure 1 demonstrates, Wayne County and other metro Detroit counties sentence to jail, probation, and prison in disparate ways compared to each other and the rest of the state. Along with these patterns, Michigan serves as an interesting site because it has shifted from mandatory to advisory sentencing guidelines, similar to the federal system and other states, as described below.
Figure 1: Percentage of dispositions to prison, jail and/or probation: comparing metro & non-metro circuits

For decades, socio-legal scholarship has recognized and devoted much research toward the problems of sentencing disparity (Frankel, 1973). In top journals, for instance, analyses of sentencing disparity are frequently published (e.g., Lynch & Omori, 2014; Ulmer & Johnson, 2017), and discussion around the impact of sentencing guidelines reforms upon disparity continues (US Sentencing Commission, 2010; Ulmer et al., 2011). Noting the ongoing yet changing use of sentencing guidelines and the continuation of disparity, understanding these sources of influence remains an important point of focus for sentencing research. Criminological explanations of sentencing disparity overwhelmingly come from quantitative analysis of official court data. Over the last several decades, these studies have offered valuable contributions through years of empirical studies and methodological improvements (Ulmer, 2012). Yet, even the most sophisticated studies leave unexplained variance. For example, an analysis by Ostrom and colleagues (2008) explained 67 percent of the variance in sentence length disparity under Michigan’s guidelines, while controlling for legal and extra-legal variables.

The current study uses in-depth interviews to build upon a small but growing set of recent qualitative studies of sentencing practices (Clair & Winter, 2016; Hester, 2017; Ulmer & Johnson, 2017) by investigating sites where disparity emanates. From drawing out perceptions from judges on their experience with sentencing, this study locates themed details around sources of disparity and contextualizes them into past empirical findings and theory. Past empirical findings provide evidence on the extent of influence by extra-legal sources of disparity, and this study offers context to how these sources create disparity.

COURTS COMMUNITIES & THE MULTI-LEVEL SOURCES OF SENTENCING DISPARITY

Researchers from a range of disciplines have substantially advanced the empirical and theoretical study of sentencing (Spohn, 2009; Ulmer, 2012). This research shows that any understanding of sentencing must consider the constellation of social forces that influence it: individual characteristics of court actors and defendants, state and federal policy, organizational
composition and resources, cultural context, and racial and economic patterns. Bushway and Forst (2013) conceptualize discretion at various levels, including how legislative discretion can impose guidelines that shape individual discretion by judges. Understanding court decision-making and overarching policy contextualizes the interactive and reciprocal relationships between levels of discretion and regulation (Ulmer & Kramer, 1998).

According to Savelberg (1992), the attempt to make the individualized sentencing process into a formal rational process is fraught with troubles. Substantive rationality, “the intrusion of economic, sociological, and ethical criteria on formal rational reasoning and decision-making,” will work in contrast to formal rules and policies, or formal rationality (Savelberg, 1992: 1347). Savelberg claims that courts develop their own “internally substantive rationales,” and in the administration of justice, they continue to resist controls upon their discretion and will persistently take into account perceived outcomes of sentencing, causes of criminal behavior, and potential social inequalities. Though controls on discretion can hem in sentencing disparity in some ways, push back from values of substantive rationality will open up room for disparity. The evidence below illustrates how sentencing disparity emanates from multiple forces: individual judges, community context, local court culture, and authoritative state institutions.

**Courts as Communities & Judges’ Decisions**

Sentencing research has largely focused on the 1) individual level, 2) the institutional and organizational level, and 3) the jurisdictional social, economic, and political context. Eisenstein, Flemming, and Nardulli (1988) brought these three branches of research together and developed the “courts as communities” theoretical perspective. By merging the multi-level influences upon court processes and outcomes, they offered a conceptual understanding of the complex nature of sentencing. Overall, Eisenstein and others (1977; 1988) argue that local court communities and workgroups develop their own formal and informal norms around case processing, and these equal or exceed the influence of state-level laws and policies toward shaping court processes and sentencing (see Ulmer & Kramer, 1998).

Largely conceptualized at the county-level, seeing courts as communities allows us to understand the long-standing sociological observation that law behaves differently across contexts (Black, 1976). Eisenstein and colleagues’ (1988: 28) clearly illustrate how county legal culture dictates many court outcomes. In their research, county legal culture “consists of the values and perceptions of the principal members of the court community about how they ought to behave and their beliefs about how they actually do behave in performing their duties.” By studying nine different courts across Pennsylvania, Michigan, and Illinois, they drew out how not only the relationships and actions of judges, prosecuting and defense attorney interact but also the influence of political and social context, other court workers, geographic and spatial dynamics, generational groupings, among other cultural and organizational factors. Status and power make a difference in how cases are resolved and whose perspective controls processes and outcomes.

Though the metaphor of communities emphasizes the power of group influence, Eisenstein and colleagues (1988) recognized that individual judges continue to hold a powerful role and exert their agency in disparate ways. Johnson (2006) empirically demonstrated that 6 percent of sentence length variation was a result of differences between Pennsylvania judges. In
this regard, a number of theoretical perspectives integrate with the “courts as communities” perspective and help us further elaborate on the complex influences of sentencing.

Evidence on how judges’ “focal concerns” drive judicial decisions blends into theory of court communities; primary concerns that dictate judicial decision-making are formed within the community based on its norms and organizational resources. The focal concerns perspective “argues that both legal and extralegal considerations affect the interpretation and prioritization of focal concerns through local substantive rationality” and that punishment decisions are determined by blameworthiness, protection of the community, and practical constraints (Ulmer, 2012: 10). Overall, we see individual-level disparities across several demographic traits such as race, age, gender, marital status, and employment (Spohn, 2009), and the focal concerns approach has tried to explain some of these differences. Steffensmeier and colleagues (1998) theorized that judges base most of their attributions regarding focal concerns on legally relevant information but also rely on attributions based upon the race, gender, or age of the defendant as they relate to the focal concerns of sentencing. Specifically, judges attribute increased blameworthiness and threat to the community toward young, black males based on racial (as well as age and gender) stereotypes. This leads to heightened sentencing penalties for being young, black, and male.

Albonetti’s (1987; 1991) research provides insights toward understanding these patterned attributions by showing how judges’ decisions are based on assumptions from limited case and offender characteristics. This perspective highlights that even though judges are making significantly impactful decisions upon individual lives and the system, they are frequently acting in a context of bounded rationality – where uncertainty regarding rehabilitation and recidivism abounds. In these weighty and publicly visible decisions, judges “attempt to manage uncertainty in the sentencing decision by developing ‘patterned responses’ that are themselves the product of an attribution process influenced by causal judgements” (Albonetti, 1991: 250). Likewise, judges have described the difficulty of evaluating cases that are “apples and oranges” and objectively assigning a blameworthiness number (Steffensmeier, Kramer & Streibel, 1993). The difficulty may lie in the range of influences beyond these three focal concerns, other relevant factors such as a defendant’s amenability to or experience with treatment, extent of substance dependency, community and family ties, and employment history (Engen et al., 2003).

While focal concerns likely matter to some extent for all judges, individual judges distinctly interpret the extent of blameworthiness, threat to public safety, and defendants’ amenability to programming such as drug courts or incarceration. Judges may interpret the threat to public safety differently for certain offenses (for example, some judges interpret drunk driving as more threatening than other judges do). They may also place different values on individual blame as opposed to structural root causes of behavior. And they may apply these attributions differently based on defendants’ identities. Further, judges hold different views of justice. Some emphasize rehabilitative, others retributive, and some restorative justice (Eisenstein et al., 1988; Belknap & McDonald, 2010). Though empirical evidence frequently theorizes sentencing disparities through the lens of focal concerns, they do not include direct measures of blameworthiness or perceived threat to the public.²

² Empirical studies use criminal history and seriousness measures, which are proxies of judges’ perceptions of blameworthiness and protection of community.
Organizational and Community Context

A core element of the court communities perspective is the organizational components, histories, and relationships within the local court (Eisenstein et al., 1988; Ulmer, 1997). In their study of Detroit courts in 1972-1973, Eisenstein and Jacob (1977) illustrate the historical emphasis on the avoidance of jail time due to legal suits and negative media attention toward jail conditions and overcrowding. Additionally, the history of severely backlogged dockets and a shortage of judges institutionalized plea bargaining to increase “efficiency” (see Eisenstein & Jacob, 1977: 126-168).

Related research focuses on organizational maintenance and efficiency as determining sentencing patterns (Dixon, 1995; Engen & Steen, 2000). Bureaucratization, local resources, informal relationships, and the extent of cooperation among courtroom actors to maintain stability in the sentencing system. “Adopting this approach, the central problem for understanding the organization of sentencing becomes defining the operational rather than political or formal legal goals of sentencing” (Dixon, 1995: 1162). Bureaucratization and larger caseloads increase the importance of plea negotiations as a way to settle cases efficiently, and these become entrenched into local court culture. Importantly, several studies show the “plea discount” or “trial penalty” - sentencing differences between cases settled by plea negotiations, jury trials, and bench trials (Ulmer & Bradley, 2006; Kramer & Ulmer, 2009; King, Soule, Steen & Weidner, 2005). The trial penalty itself varies across case traits and community contexts.

Eisenstein and others (1988) argue that court size changes these organizational features and outcomes. The size of a group affects hierarchies, role specialization, and formalization of rules and relationships. They find that large courts are more formalized and specialized, and subsequent outcomes can differ. More generally, we see that urban, suburban, and rural courts vary in their legal cultures and sentencing (Ulmer, 1997). Beyond the urban/suburban/rural divide, courts are embedded into social, economic, and political structures and histories of the community. Racial composition, socio-economic shifts, and political climate are argued to cultivate environments where these individual-level decisions (including bias and stereotyping) and organizational structures differ, and in turn, affect sentencing (Ulmer, 2012; Wang & Myers, 2015). Understanding these organizational features through research beyond state datasets continues to be a crucial site of investigation for the progression of sentencing research.

In the next section, I provide context by describing the development of guidelines in Michigan and recent examinations of the shift from mandatory to advisory in the federal system.

RESEARCH CONTEXT: MICHIGAN’S GUIDELINES & THE SHIFT FROM MANDATORY TO ADVISORY

All of these influential factors in the sentencing process occur within the context of how the state legislature defines sentencing parameters. In 1999, the Michigan State Legislature initially implemented the current sentencing guidelines as a mandatory system. However, in 2015, People v. Lockridge resulted in changing the sentencing guidelines from legislatively mandated to an advisory guideline system. Importantly, the move from mandatory to advisory guidelines opens up windows for disparity. Specifically, judges were previously required to sentence within the guidelines, and allowed to depart from these requirements only in cases with “substantial and compelling” reasons. Following Lockridge, the “substantial and compelling”
requirement no longer exists. After *Lockridge*, judges are still required to offer written reasoning and are subject to appeal following departures from the guidelines.

Briefly, sentencing guidelines attempt to restrict sentencing decisions through formal scoring processes (see Figure 2). In Michigan, offense seriousness and prior criminal history are the two scores that officially determine a sentencing grid, within which a range of limited sentence options are available to a judge. The number of possible grids is comparatively high, as each sentence is placed in 1 of 258 possible cells on the grid. There are 36 possible cells in “Class A” alone. The offense variables are scored based on factors such as the use of a weapon, extent of victim physical or psychological injury, intent of the offender, number of victims, extent of criminal sexual penetration, and number of contemporaneous felony criminal acts. The prior record variable is scored based on whether past crimes were crimes against persons, property, public order, public safety, public trust, or involved controlled substances. Additionally, habitual-offender enhancements applied (or not) via prosecutor discretion and comparatively wide sentence ranges allow for discretion and disparity (Ostrom et al., 2008).

![Figure 2: Steps to determine a sentencing guideline score in Michigan](source: Council of State Governments (2014))

---

While this current study does not directly examine the potential effects of *Lockridge*, which was a key research question for an adjoining analysis that was part of a larger project (Smith & Smith, 2019), the context of Michigan during this transforming period matters. Michigan’s change from a mandatory to an advisory system followed a number of related rulings. In two cases commonly referred to as *Booker/Fanfan*, the Supreme Court found that mandatory guidelines were not unconstitutional per se but could only be applied to the judiciary in an advisory capacity (*United States v. Booker* 543 U.S. 220, 2005). *United States v. Gall* (2007) and *Kimbrough v United States* (2007) both further enhanced judges’ discretion. These cases raise important questions about what happens when mandatory guidelines are lifted in favor of advisory guidelines (Baron-Evans & Stith, 2012; Ulmer, Light & Kramer, 2011; US Sentencing Commission, 2010).

The US Sentencing Commission conducted three evaluations post-Booker, which suggests that disparities increase over time and call for a “fix” to post-Booker disparities (USSC 2010, 2011, 2012). Scott (2010) showed evidence that in one federal district sentence lengths and inter-judge disparity both significantly increased following *Booker, Kimbrough*, and *Gall*. In direct contrast to the notion of post-Booker dynamicism, Ulmer and colleagues (2011) reanalyzed the same data as the USSC and found the idea that a policy “fix” should be called into question. They argued that there was no significant effect caused by this policy change. Baron-Evans and Stith (2012:1689) claim that after *Booker*, judges have been able to mitigate racial disparity by giving sentences below guideline recommendations. Due to the changing yet continued use of guidelines, this analysis seeks to build upon past research on the causes of disparity and provide further insight into what the “black box” of statistical analyses may reveal.

**RESEARCH DESIGN**

The quantitative study of disparities is essential toward demonstrating the scope of the problem and offering generalizable findings. Analyzing qualitative perspectives of the decision-makers is essential too. Qualitative work allows for a more nuanced, contextualized understanding of local processes, interactions, and contexts. Across many institutions, we see problematic disparity, and we must continue asking police officers why they use force differently, physicians why they provide treatment differently, bank loan officers why they give loans differently, and so on. Comparatively, few sentencing studies have asked judges about their interpretations of sentencing and disparity. This study adopts the position that examining the attitudes and beliefs of key stakeholders in the courtroom adds refinement to our collective research on the complexities that constitute sentencing and lends contextualizing evidence on how “law on the books” and “law in action” differ (Chambliss & Seidman, 1971).

These interviews seek contextual detail unseen within statistical court data. For the most part, sentencing research relies on data constructed for court purposes, not sociologically ones. Sociological important characteristics are often missed. As Ulmer (2012) and Baumer (2013) argue, sentencing research entails a continued need for increased attention upon varying details within court processes. Regression outcomes, while giving key insights, are indirect observations of the complex cultural context where sentencing is embedded. In the quantitative analysis of sentencing disparity using court data, a significant portion of the variance remains unexplained. Specification error, an elemental methodological issue in regression analyses, is an ongoing concern. This can happen in two ways: analyses have left out variables that can explain disparity, or analyses include variables that are irrelevant in explaining disparity. Qualitative
research has a long history of informing explanatory quantitative analysis, and these interviews provide such evidence.

Addressing the call for increased qualitative sentencing research, recent studies have examined the judges’ interpretations of race disparities in sentencing (Clair & Winter, 2016), the influence of judicial rotation upon court communities and sentencing uniformity (Hester, 2017), judicial perceptions of sentencing guidelines (Ulmer & Johnson, 2017), and the production of racial inequality in criminal courts (Van Cleve, 2017). This recent development builds upon past qualitative work (e.g., Eisenstein et al., 1988; Ulmer, 1997) that has complemented the more numerous quantitative analysis into courts and sentencing.

The Judges

The final sample includes 20 circuit court judges in Michigan. Circuit court judges compose a “sample of experts” – an elite and specialized group – that yields rich accounts of direct experience with sentencing. Ninety judges from twenty-three circuits were recruited by email and/or phone. This resulted in a sample of twenty judges from nine different circuits (see Table 1).

The twenty-three circuits used for recruitment were located across Michigan’s six court district regions and were selected to ensure a range of regional differences in terms of rural/suburban/urban. Sampling aimed to maximize the range of responses rather than a focus on generalization to a larger population (Clair & Winter, 2016). Sampling for range is one alternative to generalizability (Weiss 1994; Small, 2009). Therefore, while claims of generalizability cannot be made, this set of cases provides a range of diverse perspectives on sentencing among judges.

Five of the nine circuits were in metropolitan areas with large populations. Fifteen judges worked in metropolitan, and five worked in rural circuits. When a circuit was chosen, all judges within that circuit were recruited. The sample includes judges from a range of different circuits with distinct local characteristics – demographics, crime rates, caseloads and geographic differences. Years of experience with criminal sentencing at the criminal circuit court level ranged from seven months to twenty-four years; average time on the circuit court bench was eleven years among this group. Judges had differences in their past legal careers prior to taking the bench. Some reported little experience practicing criminal law before their service as a criminal sentencing judge, while others frequently worked with the sentencing guidelines. Unfortunately, we were unable to find other demographic information on all criminal circuit court judges as a way to compare the sample with the population.

Circuit court judges were the target population due to their role in all types of felony sentencing under the guidelines. There are fifty-seven circuits in Michigan. Typically, a circuit coincides directly with one county (e.g., Wayne County is the third circuit). However, in low population counties, one circuit may include two, three, or four counties. Circuit court judges are either governor-appointed or elected and are officially non-partisan appointments or elections. Elected judges serve six-year terms.

---

4 IRB protocol disallowed revealing the demographic information of individual judges based on threats to confidentiality.

5 Stephens (2015) found that 26 circuit court judges were African-American in 2010 (there are currently 237 judges in Michigan). African-American judges in this study constitute 10 percent of the sample, similar to the overall population.
Table 1: Judge Characteristics

<table>
<thead>
<tr>
<th>Jurisdiction type</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>5</td>
</tr>
<tr>
<td>Large city</td>
<td>5</td>
</tr>
<tr>
<td>Large Suburb</td>
<td>6</td>
</tr>
<tr>
<td>Medium City</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prior law experience</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Criminal defense attorney</td>
<td>6</td>
</tr>
<tr>
<td>Prosecuting attorney</td>
<td>3</td>
</tr>
<tr>
<td>Civil litigation attorney</td>
<td>6</td>
</tr>
<tr>
<td>Civil/criminal</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>2</td>
</tr>
<tr>
<td>White</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>16</td>
</tr>
<tr>
<td>Female</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Court experience</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Range of years on bench</td>
<td>0.5-24</td>
</tr>
<tr>
<td>Average years on bench</td>
<td>11</td>
</tr>
</tbody>
</table>

Interviews

This study sought a range of responses as a methodological goal in order to address its research question (Weiss, 1994). Rather than generalization, a goal was to analyze a set of responses that encompassed the range of judicial perceptions. Open-ended questions served as the primary format because they avoid the problem of overly restricting possible responses, especially when researchers are unsure about the range of possible responses (Marvasti, 2004). From this starting point, interview data provided an analytically rich portrait detailing first-hand perceptions surrounding sentencing.

The interviews were all performed face-to-face in judges’ offices from June to October 2016. The interviews averaged forty-eight minutes in length, and all but three were audio recorded. When not recorded, detailed field notes were taken during the interview. Interviews were conducted until saturation of responses was reached among a sample from a range of different jurisdiction types (i.e., rural/urban/suburban, different regions).

Defining when saturation occurs and deciding on the appropriate number of interviews are both a challenge. We, as researchers, must assess when new meaningful details around key themes have been exhausted. The original minimum goal was to perform 16 interviews distributed across different regional types (urban, rural, suburban). After I conducted 16 interviews, I revisited my notes, and wanted to further examine the theme around plea bargaining.
and Cobbs evaluations (an unanticipated issue prior to starting the interviews and a theme discussed below). An additional four interviews were conducted until saturation on main themes was reached. Following past research on the issue of sampling, I: 1) began with an expected sampling frame (16), 2) maintained flexibility until saturation was met, and 3) established a stopping criterion of two interviews when no new themes emerged (Francis et al., 2010).

Open-ended questions integrate a range of viewpoints regarding sentencing. Rather than relying on standardized surveys, semi-structured interviews allow individual respondents to discuss particularities from their own experiences (Marvasti, 2004). Since we know that sentencing disparity stretches into three basic categories - between courts, between judges, and within individual judges – allowing respondents to detail the circumstances under which courts may vary in their decisions, or when judges in the same court may vary, or when judges themselves may vary can give us guidance toward understanding sentencing.

The interviews began by asking judges about their legal background and judicial appointments. Then, open-ended questions focused on what they perceive as factors that influence the sentencing process. In their responses to questions about guidelines as well as questions specifically regarding influential factors, judges described how many complexities are involved in determining a sentence. They were asked, “there appears to be disparity in sentence lengths across judges and counties in sentencing patterns – what’s your take on this?” This question was followed by judges’ perceptions about why disparity happens. At the end of the interview, judges were asked, “is there anything else you would like to add regarding sentencing processes?” and this also elicited responses about disparity and discretion in courts.

The face-to-face nature of these interviews allowed a rapport to develop and provided opportunities to ask follow-up questions beyond those available via surveys or other data collection methods. During the beginning of the interview process, I signaled my familiarity with sentencing trends and policy in Michigan. Early on in each interview, I described my recent statistical research using Michigan data and framed the interview as a way to fill in the blanks of the statistics. Judges frequently signaled their interest in the study, hence their voluntary participation. Judges elaborated upon specific, detailed examples and showed me examples of materials such as pre-sentence investigation reports, casebooks, and local data reports that illustrated details they were describing. While field notes on my interpretation of social cues displaying openness versus guardedness were different among some judges, overall, I perceived the clear majority of discussions to be honest, thorough, and open in terms of answering the questions. Overall, judges openly responded to all questions, though with varying detail.

Analysis

Following Weiss (1994), interviews were designed to address: the core research problem (judges’ perceptions on discretion), a sense of the breadth and density of the material collected (particular sites of discretion), a repertoire of understanding founded on past research (factors previously shown to be influential in the sentencing process), allow for questions to develop across the interviews (topics that arose in early interviews were explicitly addressed in later ones), and substantive descriptions toward overall findings (detailed examples).

Data analysis followed guidelines outlined by Corbin & Strauss (2008). The initial stage of analysis involved thoroughly listening to each recording and writing memos on first impressions of the participant and their experiences. Reflections on initial interpretations of their responses were noted, and notes regarding repetitions and contrasts among participants were
tracked in memos after several were read. As Bohm (2004: p. 271) outlines, in this style of
coding “researchers use their background knowledge about the context of the textual passage
being investigated and, in general terms, their knowledge about the area of investigation” and
blend this previous knowledge with the text itself using in-vivo codes “taken directly from the
language of the field of investigation. In-vivo codes are components of ‘theories’ formulated
personally by the producers of the text in question.” Then, the coding process utilized NVivo
software to find “repetitions and regularities” and distinguish nuances in coding nodes. The
analysis examined participants’ perspectives through an open-coding process and located several
nodes (Bazeley, 2007; Corbin & Strauss, 2008). In these nodes, dimensions were conceived as a
spectrum and used comparisons to draw out extremes and contradictions (Bohm, 2004). Three
major themes were identified through the first cycle of coding, and subsequent subthemes were
located through the second cycle of coding.

FINDINGS

Judge 8 puts the core problematic of disparity into perspective by frankly stating:

I could give 15 years, and there’s nothing anyone could do about that. Completely
unfettered discretion on my part. Or I could put them on the Holmes Youthful
Training Act and give them 40 hours down at the YMCA. How does that make any
sense that two ideologically different judges could reach such disparate position?

In general, judges discussed how local jurisdictional differences, individual-level
variation among judges, and influence from other state institutions all affect sentencing disparity.
Within these three major themes are subthemes that outline mechanisms and case qualities that
judges perceive to exert patterned influence upon sentencing decisions. Many of these findings
coincide with past research (Spohn, 2009) as well as first-hand anecdotal observations by those
working in courts. Additionally, many findings elaborate upon details important within the
context of Michigan’s sentencing system.

Localization, Organizational Resources, and Local Policy

The first theme captures how judges frequently acknowledge variation across
jurisdictions. Specifically, judges describe how regional social and political contexts,
organizational resources, and policy differences matter. Differences between metro and non-
metro areas were seen as salient. Judges noted how their decisions are enabled and constrained
by organizational resources. Additionally, they decide sentences in a multi-stage justice process
that includes localized policies and norms that differ between jurisdictions.

Regional Differences. A commonly apparent observation among judges was the
difference between metro Detroit and everywhere else. The sheer volume of cases and
frequency of serious cases compared with all other counties were seen as causes of different
outcomes. More specifically, judges noted the lower sentence lengths in Wayne County (Detroit
is in Wayne County) as compared with other places. Counties outside the metro area were seen
as generally more punitive. Judge 4 explained:
I go by the assumption that large urban communities – because of the volume – we become desensitized to the level of crime and violence and as a result, you’re going to see less severe sentences maybe for particularly for certain types of crime in Wayne County than you’re going to see in a smaller community where they don’t see it, they’re not desensitized to it. And they are going to bring down the hammer.

The comparison of highly populated areas, especially metro Detroit counties, with less populated ones, was an obvious observation throughout these interviews. The simple fact that some courts see more crimes and more serious crimes than others makes a difference. While population size is related to organizational characteristics and goals, which can affect decisions, research has also shown that there is a unique independent effect of population size. Specifically, sentences in larger courts will be less punitive (Eisenstein et al., 1988), and judges with heavier violent caseloads are often more lenient with violent convictions (Johnson, 2006).

Judge 8 noted both the regional and socioeconomic factors as well as legal cultures (according to many sociological theories, legal culture is influenced by these regional and socioeconomic differences).

Two adjacent counties can be very different in demographics. So, you would expect differences due to the socioeconomic factors. Another thing is the legal culture in each is different. Do you have an aggressive prosecutor? Do you have a thoughtful prosecutor? Do you have an aggressive police presence? Do they exercise discretion a lot? Do they overlook a lot of crime?

Variation is not solely a result of individuals making their own decisions. The context matters because local community culture, courtroom culture and policy, crime patterns, among other factors, lead a place to have a certain pattern in sentencing. Explanations of jurisdictional differences, particularly in sentencing guidelines systems, illustrate that localized court norms and policies filter the formal state-level requirements and transform their use to suit local goals and resources (Ulmer & Kramer, 1998).

Organizational Goals and Resources. Within the discussion on jurisdictional differences, judges noted that availability of programming and jail space, and the larger issue of county versus state costs, guided variation in sentencing. The influence of programming and jail availability is a less-examined focus in sentencing research though studies show the effect of available jail space on sentencing (Ulmer & Johnson, 2004; Johnson, 2006; Kramer & Ulmer, 2009). As Ulmer (1997) describes, court actors often dislike the fact that jail space is considered in sentencing, yet they admit there is little they can do about it.

In terms of programming, drug courts were the most frequently available for these judges; several judges also indicated that veterans court, mental health court, treatment programs, and sobriety court were among programs used. Judge 2 discussed the different availability of programming:

Counties like ours that have long recognized the value of alternative programs... I mean, let’s be honest, [addiction] is what brings people here. And to focus on rehab and as opposed to just punishment...this gives us flexibility to continue to do that without gaming the system. Whereas other places around the state, judges tend to
be more political, more aligned with law enforcement, and ‘you’re not a good judge if you’re not sending a third or half of your felons to jail.’

Judge 3 contributes a similar response about programming in the context of discussing the merits and drawbacks of guidelines, “In one sense you want some predictability, some consistency. On the other hand, sitting in [this] county, we have far more resources on the local level than a lot of counties in the state, so we have options [besides prison].”

In these two examples, judges were from the same county, and they discussed how their county had a residential probation program that offered a number of on-site services. Additionally, this county had a veteran’s court, mental health court, “swift & sure” probation, and sobriety court. All of which are not offered everywhere. Judge 8 stated that, for example, veterans’ courts are not available in some smaller counties because there are simply not enough defendants who are veterans. Lack of availability could be reasonably assumed to be influential in rural/urban sentencing differences. Judge 11 discusses similar experiences and the perceived effectiveness of programming:

I’ve been really lucky to have people around me who are pretty visionary….our community corrections director is very knowledgeable and has put together an outpatient program that is all aimed at anti-social thinking and behavior.

Perceived effectiveness, types of programs, and program management varied across circuits and judges. Thus, when applying these findings to quantitative analysis, the mere availability of programming should not be measured as universally impactful on sentencing across places. First, the use of programs depends on how judges perceive them as effective. Judges from one circuit detailed how their circuit had two different drug courts managed by two judges with extremely different views on punishment. In terms of management differences, some counties have a single judge that works with all drug court cases. In other places, multiple judges oversee drug court. Judge 9 questioned the influence of programming in their own sentencing decisions:

We have a drug court here, and that includes alcohol. I think addiction is conceivably the greatest social problem we have in our communities, and I think it’s really, really important to devote money and time to it. I just don’t think the court is the place to do it. I really have a conflict with being in law enforcement, then at the same time saying ‘well, if you participate in rehab and I check you every week maybe you can get out of that offense.’ I’ve got a problem with it. I know they’re somewhat effective…but the court shouldn’t be doing that.

These more reluctant responses toward specialty courts and other programming illustrate how there is likely to be multiple orientations and uses of these options, even when available.

Connected to program availability is jail capacity within the circuit. Judge 9 claimed that they were unable to sentence to jail sometimes because the jail was filled with people convicted of less serious offenses from the district court. Judge 18 stated that along with specialty courts, “jail space fuels a lot” of the sentencing decisions. Judge 8 described the importance of jail,
One of the factors is that jail capacity is there. It’s an available tool...I bet that is true across the state, with crime rates going down, jail space is more available. So, you can take people that you might not otherwise, and sentence them to jail.

As many past observers have noted, this judge also detailed that available space can bring pressure to put people in local beds, whereas the opposite of no space can be an influence in some jurisdictions. Depending on the conviction, no jail space could increase the likelihood of probation or a prison sentence.

These first-hand accounts of working judges suggest that empirical analyses of sentencing must include controls for jail space, programming, and state-level financial budgets. Judge 16 stated simply as a matter of money, “we can’t put everyone in jail or prison.” Further, in regards to financial drivers of sentencing decisions, Judges 5, 6, and 8 noted the financial reality that a prison sentence turns the immediate financial responsibility over to the state as opposed to the county. Prisons cost the state money; jails cost counties money. There was description of this back-and-forth relationship between the state and counties. Some judges noted how the state pushes for county time while there is a simultaneous incentive for some counties not to offer county-funded programs and, in turn, send more to state prisons on the “state dime.” Judge 19 argued that MDOC recommends non-prison sentences – via presentence investigation reports – due to financial burdens.

They’ll do anything to not get someone into prison. If it’s a straddle cell,⁶ they’re almost always going to recommend local jail…there’s this policy that they’re trying to avoid prison. [But] I look at it as individual justice. I don’t care about economic consequence... we’re at loggerheads.

These differences in resources across jurisdictions raise theoretically important reasons why we see disparity. Not only do alternative sentences to programming open “windows of discretion,” but because they include priorities of rehabilitation, they mandate consideration of substantive characteristics of the cases that drive disparities (Engen et al., 2003). Additionally, community- and state-level social and economic forces influence the availability of alternatives.

Policy & Plea Agreements: Different Court Plea Norms. Emerging from these twenty judges’ responses was the identification of court policy and norm differences between circuits in plea agreements. These differences have consequences; much research and first-hand experiences tell us that convictions by plea result in significantly lesser sentences compared with those of a bench or jury trial (King et al., 2005). Judge 20 describes an example of how court policy is shaped internally and can define a much different negotiation process compared with other circuits. In this county, judges were put in the position to play more of a role in plea negotiations as a response to a “no pleas for defendants with priors” policy. This becomes important considering the influence of priors and plea convictions upon sentencing.

The negotiation used to happen between defense attorney and prosecuting attorney. Then [the prosecutors’ office made new policy] that rank and file prosecuting attorneys can’t offer a defendant anything [in terms of plea deals] if they have any

---

⁶ Straddle cells are those in which the lower limit of the recommended range is one year or less and the upper limit of the range is more the 18 months. When a guideline range falls within a straddle cell, judges can decide to give a prison sentence or an intermediate sanction, including jail time.
prior felony...so then what happens is it pushes everything to the judges, and that’s why I think some judges don’t want to [negotiate pleas]. Because it’s like ‘I shouldn’t be doing this.’ And if the PA wants to plea and the defendant wants to plea, and they are looking for some guarantee of something, instead of the plea bargain of ‘I’ll plea to count 1, dismiss count 2, its ‘well I’ll plea to count 1 and plea to count 2,’ and the sentencing guidelines are 7 months to whatever…’would you agree to county time as opposed to [prison] time?’ Or I’ll cap it at bottom third of the guidelines, and the attorney can still argue for a lower sentence. [T]he judge takes on more of a role of the plea bargain…where it really should be between the defense and the PA.

Judge 2 outlined processes in plea deals in a county with a “no-charge bargaining” policy:

You would just develop workarounds. Even though this a Class D offense…if you got the right Assistant Prosecutor and you’re the defense, you’d come in and agree to score this on the G grid. That was fairly routine especially in counties like this because they have a no-charge bargaining policy…That was very common to either stipulate to the scoring of the guidelines…They would [say], ‘Judge, we are going to plea to the charge, we are going to recommend probation, and we stipulate that scoring of the guidelines is an A1.’ A lot of that happened. So, you’re not truly scoring the guidelines...You’re stipulating to it and there’s lots of winks and nods.

Judge 8 added, “Our prosecutor has certain classes of crime: they don’t plead second marijuana charges, they don’t plead failure to pay child support, OUIL third…those handful of cases our local lawyers know they’re just not going to get a plea bargain.”

Court policies on charge and plea bargaining and the application of habitual sentence enhancements are variables commonly missed by quantitative court data. Here, judges identified different court-specific policies, which in turn may result in different sentencing patterns for different types of offenders and the varying use of plea negotiations versus trials. Similar to policy on pleas, judges explained that there were court-specific policies set by the prosecutor’s office regarding habitual offenders. In at least one court in this sample, the prosecutor’s office mandated that all defendants eligible for the habitual enhancement receive it. Most courts selectively apply it.

When discussing sentencing guidelines and what mechanics are important toward understanding the sentencing process, twelve judges talked about Cobbs evaluations.\footnote{A Cobbs evaluation (often called a Cobbs agreement) basically allows a defendant to evaluate what the possible sentence would be from the judge, in turn gaining some certainty about the sentence before pleading guilty. The judge does not need to follow the “agreement” however. A Cobbs evaluation does allow the defendant to withdraw the plea. (\textit{People v. Cobbs} 1993).}

The Cobbs evaluation is unique to Michigan. Judges often described the important difference in judges who use them and those who do not. Typically, this was described as a judge-by-judge decision. But there were mentions of how it was within a court’s norms – though not official policy - to not use Cobbs evaluations. Adding to the discussion of judicial involvement in plea agreements, Judge 20 said:
I do a lot of Cobbs agreements. Most of the judges in the building do them... For us, Cobbs are essential [to process cases] because we have a situation here, over the years, there’s been less and less pleas. And from an attorney’s perspective, if your client is going to plea, you want to show you’re at least doing something for them... get rid of two counts or something.

Judge 15 spoke of the advantages of Cobbs evaluations: they stated that defendants know the “cap” or maximum sentence that is likely to be imposed before they plea. This judge estimated that 75 percent of pleas go with the Cobbs evaluation, and in turn, save the court time because negotiation of the sentence is less time-consuming and a trial is avoided. On the other hand, judges note the disadvantages to Cobbs evaluations during plea negotiations and how they are administered differently across place. Judge 10 argued that Cobbs evaluations turned justice into “commerce,” while others argued that the role of the judge should not include participation in plea negotiations. Judge 12 said, “I have never given a Cobbs evaluation because I think it’s an invasion. It’s something a judge shouldn’t engage in... never would, never will.”

Here, judges outwardly express their use as a means to the end of getting cases completed and simultaneously detail how Cobbs evaluations contradict larger goals of justice. At the same time, disparity between similar offenses can emerge from differential use of Cobbs evaluations and plea deals.

Policy & Plea Agreements: Explaining and Justifying Different Outcomes for Pleas and Trials. These differences in practice matter due to the disparate sentences between convictions by plea and trial. As much empirical research has illustrated (Ulmer & Bradley, 2006), the decision to exercise one’s constitutional right to a trial or be disallowed by the prosecutor to enter a plea agreement leads to much higher penalties. Judges were asked their perspectives on disparate sentences between convictions by plea and trial. Many findings mirror those from past research (Ulmer & Kramer, 1996; Ulmer, 1997). Judges acknowledge that pleas save courts time, that exhibiting remorse through a guilty plea is important, and that the troubling details found during trial increase blameworthiness. They highlighted particularly how “ugly facts” in severe violent cases matter, which supports evidence that the penalty for going to trial is especially salient in more serious offenses (Ulmer & Bradley, 2006). Judge 14 explained:

I guess there’s probably a philosophy among judges: ‘if you save the victim time, and you save the state money and you don’t make me listen to what you did, I may be willing to go to the bottom of the sentencing guidelines. But if you put everyone through this, including the victim, and drag people in here, cost the state money, or whatever, then that might not be available.’

Judge 15 said that during trials “you hear additional things usually not favorable for the defendant,” especially in criminal sexual conduct cases. Not making victims go through it all again is something they consider when sentencing plea convictions in CSC cases and assaultive cases; pleas may have reduced time because they avoid traumatic experience for victims. Likewise, Judge 16 said, “in trial you hear gory details” and when the “conscience of the community decides the guilt there is a higher likely to sentence at the upper end of the guidelines.” Judge 9 noted a similar dynamic:
If they go to trial, it’s not that I make an effort to do something different but when you sit through a trial about violence or armed robbery and you hear the details and see maybe whether someone’s lying or being a jerk, it can affect how you view the case.

Judge 19 stated they were much more likely to sentence on the upper end of the sentencing guidelines range or depart upward after a trial conviction because “When you go through the trial, and you have a victim on the stand, or you learn more about the defendant than what a cold pre-sentence investigation would give you. You have more ammunition, so to speak, for a harsher penalty.” As compared to a plea agreement where a pre-sentence investigation is used:

…then you go to trial and you find out the weapon is super large, maybe an assault weapon – nobody told you that. The victim is on the stand disassembling, crying, talking about how they had nightmares, in absolute panic, thought they were going to die, going through all that mental anguish, and you have a defendant who is particularly belligerent and obviously lying. Then that’s a different set of circumstances.

Empirical analyses have demonstrated the plea/trial disparity, and it is often interpreted – as it was amongst many of these judges - as a “necessary evil” among many court officials. The experiences of judges occur in a context where courts are overburdened with caseloads that prevent most cases from being decided by trial. The necessity of pleas is driven mainly by society’s reliance on formalized criminal courts to address a range of social problems without the resources to consistently use jury trials (Alschuler, 1979; Reiman & Leighton, 2016).

Another factor in the plea/trial disparity is the influence of previous stages: charge bargaining, cooperation with law enforcement, and the intersection of counting priors and/or habitual enhancements. Judge 16 discussed how an upward departure is sometimes a result of a plea to lesser charges (CSC-3 rather than CSC-1) but the agreed upon sentence is higher than what guidelines say is the CSC-3 charge, but less than CSC-1. In this case, a lower sentence appears quantitatively as a more punitive sentence. Judge 7 pointed out that the intersection of habitual sentencing enhancements and plea deals:

It’s like ‘we’ll drop your status as fourth habitual offender if you plead to the underlying charge.’ So, I find [habitual enhancements] effective in getting sentences resolved, knowing they have that potential penalty. The only time it would generally be imposed is if they go to trial, so it’s kind of a nice negotiating tool.

In these cases, the definition of habitual offenders becomes blurred by negotiations. Not all those eligible habitual offender enhancements are applied (Crawford, Chiricos & Kleck, 1998). From judicial descriptions, habituals were more likely to be applied for trial convictions and more likely dropped for pleas. In terms of quantitative analyses, this shows the difficulty of defining these key variables and comparing like cases, and illustrates the need for the inclusion of multiple court stages in robust sentencing studies. On top of this methodological issue, responses throughout the interviews, such as Judge 7 above, outline the nuanced mechanisms that allow courts to persuade defendants to plea agreements and, in turn speedily complete
convictions. Organizational needs reproduce disparity (and resist uniformity), showing how guidelines face many obstacles in the attempt to produce uniformity.

**Individual Differences: Judicial Philosophy & Arbitrary Definitions of Focal Concerns**

Judges also unsurprisingly argued that sentencing is very much an individual process, based on judicial beliefs and the role of a judge. These were not necessarily mutually exclusive explanations as many judges talked about how community, organization, and individual factors all converged to guide sentencing patterns. This theme builds upon research on the interactions between judicial philosophy of punishment, focal concerns, guidelines, and the emotional and mental difficulty of deciding sentences in a context of uncertainty and risk (Ulmer, 1997; Steffensmeier, Ulmer & Kramer, 1998). Additionally, they offer insight into how individual actors’ discretion interacts with discretion made by lawmakers (Bushway & Forst, 2013).

*Judges’ Philosophy of Justice.* Judges recognized that sentencing remains a process with human agency. Though judges practice in a bureaucratized space with rules dictating many steps, the human element of individual agency demonstrates the power the position still holds, which becomes more pronounced once guidelines are loosened (Ulmer & Kramer, 1998). Indeed, seven judges affirmatively stated, without direct questions about the issue, they have “a lot of power.” And this position, held by someone whose decisions are guided by their personal experience – whether that same day or throughout their life-course – contains arbitrariness and “the power to save or destroy lives” (Judge 11). Judge 6 stated:

> I am imposing sort of my view, and whatever shaped me for 50-some years…the same [person] from [another county]…inevitably their view is going to be different than mine. But we could have the same case, same circumstance, same crime…the sentences could be different.

Many judges were also quick to point out that within the same court, there were different worldviews on justice and distinct sentencing habits among counterparts in the same building. Indeed, even within the same court community - that holds the same protocols, electorate population, and organizational needs – there is disparity based on a judge’s individual decisions (see Johnson, 2006). For example, discussing how there are differences within the use of sentencing guidelines, Judge 19 argued:

> From a practical side, it’s much easier to have sentencing guidelines, you’re not trying to figure out what merits a just sent… it also does provide more consistency... On the other hand, you still have different judicial philosophies…and I’m considered to be one of the harder sentencers. If you had the same person before me …chances are your sentences are going to be very different. [A judge down the hall] feels like they’re a social worker, I don’t think I’m a social worker...we have different philosophies...We’ve both been elected by ‘the people’ [in the same jurisdiction].

Judge 8 also talked about varying philosophies within the same court building, where “one person who is far-right nut job, and somebody else who’s way off on the left and they’re arguing all the time.” And how some courts work “collegially and collaboratively” while others have
“no communication.” Though conclusions from past research, and within these judicial perspectives (Smith & Smith, 2019), show how rationalized justice increasingly stamps out individual agency (Heydebrand & Seron, 1990; Feeley & Simon, 1992) there are also spaces where individualized judicial interpretations can be exercised.

Arbitrary Definitions in Determining Factors in Sentencing. As in past research (Steffensmeier et al. 1998), judges in these interviews discussed the importance of focal concerns in decision-making – trying to assess the case seriousness and to what extent should prior life histories be mitigating or aggravating factors. They also highlighted the difficulty of sentencing due to unclear definitions of culpability and uncertainty regarding recidivism. While statistical research shows the importance of blameworthiness and public safety, judges’ direct accounts add detail to how blame and public threats are constructed. In discussions of individual traits that shaped decisions, judges identified many factors: age, addiction, past rehabilitation, employment, race and ethnicity, family background for young adults, mental illness, housing, children to support, family, and community support systems. Judges justified the inclusion of some factors (e.g., histories of probation and programming, demeanor in court) while others are acknowledged but viewed as problematic (e.g., plea bargaining, race).

Even given the facts of the case, judges explained how scoring guidelines are still arbitrary. The final score on the two guideline variables - prior record and offense seriousness - places a case in a cell on the guidelines grid. In this process, there are two fundamental locations for variation. First, in the scoring (e.g., assessing the extent of victim injury), and second in the assignment of a sentence given the options of a particular cell on the grid (e.g., sentencing to the lowest minimum sentence in the cell or the highest minimum sentence in the cell). Judge 17 illustrates the difficulty of defining cases:

When you sit down, and your job on a daily basis is to decide how bad murder is, how bad rape is. And where you’re going to sentence, what you’re going to do with somebody...it’s kind of a crazy business. Because a rape isn’t a rape, a robbery isn’t a robbery, a jaywalking isn’t a jaywalking. I don’t care what kind of crime you take. You’ve got to look at the mitigating factors, the aggravating factors, the background of the individuals, the safety of the community. I mean, you’ve got a lot of balls in the air you’re juggling and trying to balance out.

Decisions on where to sentence within a guideline grid and how to calculate the guidelines score are rooted in a system where both uniformity and disparity can be problematic (Baron-Evans & Stith, 2012). Judge 11 drew out some of the complexities:

Here’s the problem: I think that clearly case-by-case there should be huge disparity. There’s a big difference in the same crime. Again, it could have to do with your age, your mental health. I don’t think that everyone who does the same crime should get the same sentence. What bothers me more though is how it’s totally luck of the draw which judge you land in front of.

---

8 As an example, in Sentencing Grid for Class A offenses, cell IA has a lowest minimum sentence of 21 months and a highest minimum sentence of 35 months for a non-habitual conviction. See the Michigan Sentencing Guidelines Manual at: www.michigan.gov/corrections/
Judge 11, along with others, advocated the use of scientific risk assessments in order to systematize sentencing decisions based on recidivism research. Judges often admitted that sentencing decisions have no scientific formula. While several situational factors are considered beyond those outlined in guidelines, sentencing is unguided by empirical research on what reduces recidivism. Likewise, according to many in this group, judges do not receive training in this regard. Their expertise lies in the administration of law, not criminological risk factors. Judge 8 described this arbitrariness, “Age, mental illness, developmental phases that a person might be in and just a whole bunch of factors, most of which are my armchair psychology.”

And, continued regarding arbitrariness and a lack of training:

I get back to the training of judges. We’re elected. All we need is a law degree. We go to one week of new judges’ school, where essentially, we are taught how to zip up our robe, and roll around our chair without rolling on our robe. [Learn] court rules and basic rules of evidence, and go to it. And I can take 5 years of their life. I can just take it because I feel like it. Or, I could not. That doesn’t make any sense to me. I mean, really? I have that authority but I can wield it without a demonstrable rational basis. I can wield it on the basis of ideology. It’s a crazy system.

These accounts of uncertainty contextualize the theoretical understanding of racial disparity. Albonetti (1991) argues that problematic disparities occur in this context of uncertainty because “decision-makers fall back on attributions about reoffending risk and/or rehabilitation potential that can be linked to race and gender, and other social status stereotypes.” Theoretically, a stable and enduring criminal disposition is attributed to racially marginalized groups, while a temporary or situational involvement in crime is attributed to whites (Bridges & Steen, 1998). These racialized attributions “provide a basis for arriving at a rational decision in a domain of responsibility characterized by uncertainty,” where “[a]ttributions of a stable and enduring disposition are expected to increase sentence severity. Attributes of a temporary or situational involvement in crime are expected to decrease sentence severity” (Albonetti, 1991: 250).

As part of the discussion of problems in sentencing disparity, judges frequently acknowledged racial inequalities in the system. For example, Judge 8 stated that their court “tries to be race-neutral, but we’re probably not.” Judge 4 argued that sentencing guidelines curtail race disparity though it may still permeate into overall patterns. Judge 5 also noted how judges’ personal experiences, particularly as it relates to race, have influence. They describe that “judges will draw from their experiences on evaluating whether a person deserves a particular sentence.” And these experiences among judges are “very similar…all white communities.” Though race disparity was commonly recognized among judges, Judge 5 drew out a subtheme noted only by three of the judges; that judicial decisions are embedded in a social situation where judges are disconnected from the defendants’ communities because they are mostly white and living in white communities. In turn, as Judge 5 argued, this affects court decisions because of a severe lack of familiarity with the life circumstances of many defendants.

People try to do the best that they can. They just don’t appreciate the same things. If I’m a Detroiter and I get stopped by the police three times, a judge could look and say ‘you must be doing something wrong’ because they’re comparing me to someone in Livonia that has never been stopped by the police. And it’s like ‘hold it’ three times in Detroit equals zero times in Grosse Pointe. Three times for a black
man, zero times for a white man. Those are equal. But if you don’t think that people get treated differently because they’re black by the police in Detroit then that’s not an equation that you think has any merit.

Judge 11 argued that “The New Jim Crow” plays out in court, and that the system is loaded with inconsistencies and inequalities, and demonstrated a critique of the “get tough” sentencing pattern:

I have a lot of colleagues here; I’m just amazed…it’s like the carpenter with the hammer. Every problem is a nail. That’s the only instrument they’ve ever learned about incarceration. Like I say, I fight every day to see the human being in front me and it’s not easy when they’re brought up there on a chain and sit in the jury box and you have about 4 or 5 minutes to adjudicate the case.

Three judges argued that guidelines’ overall punitiveness heightened race disparities on an aggregate-level, and five others argued that guidelines helped to rein in individually biased judicial decisions. While a small number of judges, exemplified in statements by Judge 5, noted the importance of the “white racial frame” (Feagin, 2013), racial attributions (Bridges & Steen, 1998), and disparities, most judges did not mention or go in-depth into the role of race and court processes. While there was a recognition that race disparities exist, only a select number of judges articulated upon systemic or implicit bias.

Much research continues to be dedicated toward explaining racial disparities in sentencing and the justice system overall. Judges’ accounts point toward the calculation of guidelines scores as a source of disparity due to its arbitrariness. These accounts support the observation of ways race signals differing degrees of blameworthiness and criminal risk within the challenging human process of assigning numerical values to acts of crime (Albonetti, 1991; Bridges & Steen, 1998).

**Pressure from Outside and Above**

Another theme emerged that outlined how the Michigan Department of Corrections and Court of Appeals influence sentencing. While judges exercise discretion and hold much power, their decisions are often bounded by recommendations from the MDOC and appeals to the appellate court.

*Micigan Department of Corrections & Pre-sentence Investigation Reports.* Judges noted the importance of the pre-sentence investigation report (PSI) from the probation department and stated they relied at least partly upon its recommendations. Some said they relied on them heavily; others mentioned that they reviewed them but did not place much stock in their recommendations. Judge 4 stated, “I try to rely on my probation officers who I hope have a whole lot more training than I do about what type of program we should get for people.” Similarly, Judge 1 described a good relationship with the probation office in writing the PSI and that the “Michigan Department of Corrections does an excellent job of putting together their recommendations.” And that judges “meet with [probation officers writing PSIs] too and address things that they might be concerned about.” This judge estimated that disparity “might depend on the county too, and who’s writing the PSIs.”

---

Judge 15 stated that the PSI was a “starting point” but also likes to listen to the different sides of the case and evaluate the defendant’s demeanor, remorse, and attitude within the courtroom. Judge 8 further explained the anchoring influence that PSIs have and how it may be a source of disparity.

There is a tendency for all parties to get tethered to the number that the MDOC pre-sentence writer puts in there… If they recommend that this person receives 6 months in county jail, people are immediately tethered to that number, and as judge I am too. I say ‘oh, well that sounds good.’ The defendant says ‘yeah, that sounds good.’ Prosecutor looks and says ‘well, it could be a little different but that’s ok.’ And that is our starting point. The difference that I’ve found is that the different PSI writers have dramatically different recommendations. Some are very jail happy, while others are ridiculously probation happy.

Critique of PSIs was not absent. Judge 11 described problems:

The crap that goes into police reports…they’re writing whatever the hell….and to have us use that, as cold hard facts?...In a pre-sentence report there’s a thing called the agent’s description of the offense, and what the probation agents do is that they lift, word-for-word, the language out of the police report. It has nothing to do with what was proven in the trial; what was admitted by the defendant. It repeats all the crap that we’ve been told to not take into consideration.

A continuing body of research shows the importance of studying court processes as the multi-stage system that it is. Sentencing is not the result of one decision by one judge, but instead the end result of a process involving factors well beyond the judge (Spohn, 2009; Ulmer, 2012). The inclusion of pre-sentence reports is one piece that is frequently missing in large sentencing data sets, and yet one exerting important and varying influence according to these judges.

State Court of Appeals. The influence of the Court of Appeals was noted among judges as well. In consideration of contextual factors in the examination of sentencing, Ulmer and Johnson (2017) found that “coercive constraint” from the Court of Appeals could partly shape whether judges depart from guidelines. That is if judges are relatively certain their sentence will be successfully appealed, it limits departures. Judge 11 described the influence of the Court of Appeals and argued they would be more likely to depart downward and offer more alternatives to incarceration if this constraint was absent, “I would sentence far differently than I do, which is I think considered to be fairly lenient if I weren’t worried about appeals.”

One judge described how their large, urban court had a section in the prosecutor’s office that focused only on appellate cases. When asked how often downward departures were appealed, Judge 5 replied, “Almost all of them when you go beneath the guidelines. The prosecution has an appellate section that just does that. They just work to appeal. It’s built in. It’s what they do.” The “coercive constraint” appellate courts exert upon sentencing options, in combination with the likelihood of prosecutorial appeals, is yet another variable missed by analyses of court data.
DISCUSSION

This collection of judicial perspectives has implications toward several of the theoretical understandings of sentencing from the last few decades of research (Ulmer, 2012). Here, I focus on two core implications. First, these findings demonstrate the multiple, interacting sources of substantive rationality leading to disparity. Evidence illustrates how substantive rationality makes the goal of consistency under guidelines an elusive one (Savelsberg, 1992). These judges perceive a complex constellation of factors that influence sentencing beyond the two guidelines scores: offense seriousness and prior criminal record. Specifically, judges describe how local norms, policies, and resources of court communities, which are structured by larger institutional arrangements, guide judicial decision-making. Judges define focal concerns of blameworthiness, public safety, and practical constraints in these contextual particularities. These observations support and complicate past studies of sentencing, and show complexity in the larger empirical endeavor of understanding sentencing disparity and inequality.

Second, and building on the first implication, findings show the methodological and theoretical importance of examining sentencing as a multi-stage and multi-layered, interactive process. This goes beyond the now-common multi-level analysis of community and individual-level factors. Judges’ descriptions outline the interactive relationships between state sentencing policy, state institutions (Supreme Court, Court of Appeals, Department of Corrections), local court culture and policy, local community culture and population, and individual background and philosophies of judges.

None of these factors is independent of the others, and any comprehensive attempt to understand sentencing disparity must examine these interacting influences. Judges frequently described how a community’s context – crime rates, social demographics, politics - influences local policy and resources. These affect the range of options and norms that guide judicial decisions based on focal concerns and attributions of culpability and amenability toward rehabilitation. All of these are embedded in a guidelines system that was created by the legislature and altered by the Michigan Supreme Court. Lastly, even under the structural, legal, and normative constraints, judges have windows (of varying sizes) to exercise their own philosophy of justice.

The general perspective that local organizational structure, resources, and policy matter support Ulmer’s (1997: 189) analysis that, “the nature and character of justice and formal social control…depend as much or more on the processual orders of local courts as they do on the policies and laws of larger-scale state actors.” And, as Lynch and Omori (2014: 412) describe, a state’s criminal court system is a “semi-autonomous set of systems governed by the same formal rules, statutes, and procedural policies, while also embedded in localized legal cultures that are themselves shaped by regionally specific historical contingencies and norms.”

The common observation among these judges that metro Detroit has significantly different sentencing patterns, particularly in Wayne County where Detroit is located, illustrates the substantive local and organizational interests that resist the goal of uniformity. According to judges, the regularity of seeing serious cases and the sheer size of a court circuit matter. As we see size increase, we see corresponding changes in values, roles, resource sharing, and standardization of rules (Eisenstein et al., 1988). As judges remarked, in larger courts, some offenses are more likely to be viewed as routine, and a “going rate” was more embedded, and in turn, lead to lower or less punitive sentences, at least for certain types of crime. The availability
of local resources – jail space, drug/sobriety court, veterans court, or mental health court -
determined sentencing options.

Additionally, plea bargains, a long-critiqued practice, introduce many complexities into
understanding the fine-tuned ways that disparities persist in court processes. Judges notice the
plea-trial disparity and explain how the use of plea negotiations varies due to court policies as
well as judges’ individual approach to pleas. Ulmer (1997) observed that even under guidelines,
judges maintain discretion through a balance of power within a court community. Notably,
judges determine whether a Cobbs evaluation is used, which plays a determinant role in plea
agreements. This stage of discretion – to use a Cobbs evaluation or not – is one more hidden
stage in the multi-stage court process that often ends with disparities. As judges describe, pleas
complicate not only the production of equality and consistency in a meaningful way but also
complicate the ability to empirically evaluate the equality of sentences.

Before discussing the implications for future research and policy, it is important to note
this study’s limitations. One is that prosecutors, probation officers who draft
sentencing reports, defense attorneys, and other personnel were not included. Second, twenty judges do not
represent all criminal circuit court judges across the state. There may be some important
perceptions missed here. Third, as humans, we are not always that good at recalling our past
actions or predicting our future behavior. Any interpretation or application of these findings must
keep in mind that perceptions among judges are just that: perceptions. Themes do not
necessarily reflect empirical outcomes. In the discussion of what influences sentences, judges’
perceptions could vary from their actual past sentence decisions. Noting these limitations and
keeping in mind that this unique sample provided rich data, these detailed findings contribute to
our collective effort to understand how judges navigate the administration of justice within a
system of sentencing guidelines.

A resounding implication is the disparate use of alternatives to incarceration. Related past
research (Smith & Smith, 2019) notes that Michigan judges do not expect substantial change in
sentencing following the shift to advisory guidelines. This, along with the current findings,
provides context to important questions about the effect of Lockridge on decarceration and
disparity. The availability and reliance on alternative programming have substantial influence on
disparity and how we address problematic responses to crime. Jurisdictions now have renewed
opportunity to form these alternatives.

Verma (2015) argues that the “law-before,”” or past organizational practices and power
arrangements, influences the implementation of decarceration policy. Applying this same logic,
variation in the use of sentencing guidelines is tied into a court’s history of using incarceration
and alternatives. The temporal context matters in terms of a court’s sentencing patterns (Ulmer
1995). Court actors’ “shared pasts” can also matter; robust “shared pasts” via interactions and
negotiations increases stability and familiarity. These pasts can influence how local courts react
to alternatives to incarceration – whether they continue upon past trends or are open to
alternatives. Future analyses could examine why some courts have alternatives programming
while others do not. As incarceration rates disparately fall across the US, future research should
examine what causes some states and counties to lower their use of prison and jail while others
do not.
REFERENCES


Cases Cited:

**Justin Smith** is an Assistant Professor in the Department of Sociology and Criminology at UNCW. His research focuses on racial inequality, particularly as it relates to prisons, courts, and the war on drugs. A current project examines social activism among formerly incarcerated people. He is an instructor in the Inside Out Prison Exchange Program which involves university students and incarcerated students taking a sociology class together.