ABSTRACT

Title of Dissertation: MINDING THE GAP: AN EVALUATION OF THE POTENTIAL IMPACT OF EVIDENCE-BASED SENTENCING ON SOCIAL INEQUALITY

Rebecca Richardson, Doctor of Philosophy, 2018

Dissertation directed by: Professor Brian D. Johnson, Department of Criminology and Criminal Justice

Social inequality has been a popular topic of inquiry in the criminal sentencing literature for decades, but the effects of innovations like evidence-based sentencing on inequality have not been investigated. Evidence-based sentencing, the use of actuarial assessments to inform sentencing decisions, is a data-driven approach to sentencing that highlights public safety, promotes more selective and effective use of incarceration, and may foster both transparency and objectivity in punishment decisions. However, the practice is not without its critics, and one of the more prominent criticisms of evidence-based sentencing is that it will worsen social inequality in sentencing. This dissertation uses simulation procedures and a unique dataset that combines official court records from Connecticut with results from Level of Service Inventory-Revised (LSI-R) risk assessments to inform that concern.

An assessment of disparities in Connecticut's current sentencing system reveals disparities according to race, ethnicity, gender, and socioeconomic status (SES) that cannot be fully explained by legal and case processing characteristics or by risk factors drawn directly from the LSI-R. Several risk factors, most notably those related to offenders' residential situation, companions, mental health, and attitudes toward convention and the criminal justice system, also have their own direct influences on punishment. Disparities in LSI-R composite and domain scores are observable as well. While similar composite scores mask substantial variation across domain scores for racial, ethnic, and gender groups, low-SES offenders receive higher scores in every domain and in the composite, regardless of which SES indicator is considered. A simulation procedure further shows that disparities in the LSI-R have the potential to translate into disparities in punishment that exceed those already present in Connecticut, particularly in the decision to incarcerate.

This dissertation suggests that evidence-based sentencing may come at a social cost. It underscores the need for more empirical research on evidence-based sentencing and sources of sociodemographic inequality in punishment. It also invites discussion about the treatment of low-SES offenders in the criminal justice system, about whether actuarial risk assessments designed for correctional settings can and should be adapted to inform sentencing decisions, and about the tenuous balance between effectiveness and fairness in sentencing.

MINDING THE GAP:
AN EVALUATION OF THE POTENTIAL IMPACT OF EVIDENCE-BASED SENTENCING ON SOCIAL INEQUALITY

by

Rebecca Richardson

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:

Professor Brian Johnson, Chair
Professor James Lynch
Professor Laura Dugan
Assistant Professor Lauren Porter
Associate Professor Patrick Wohlfarth, Dean's Representative

Acknowledgements

It took a village, and though I will never be able to fully express my gratitude, I'll do my best. Thank you first to Dr. Brian Johnson, who served as my dissertation chair and has been my advisor and mentor through six years of graduate school. His academic guidance, his perspective on the field of criminology, and (most of) his practical advice have helped me become a much better scholar. I must also thank my committee members for their time and their thoughtful, constructive feedback on my dissertation.

Next I want to acknowledge the Connecticut Sentencing Commission, whose vision and commitment to improving the criminal justice system made this project possible. Special thanks are also owed to the Connecticut Judicial Branch Court Support Services Division, who compiled and taught me about their administrative data.

My family, friends, and colleagues never failed to encourage and support me as I took and taught classes, wrote a thesis, studied for comps, and finally wrote a dissertation. It was a lot of work, but they made it bearable and sometimes even fun! I am very grateful.

And finally, four years ago I met a guy who has since endured countless rounds of "Sorry but I have to write tonight" and "I don't have enough money for that". Mike, you are the bacon to my pancakes. I don't know what I would have done without your love and support. I thank my lucky stars that you showed up and stuck around.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

The investigation of social inequality has long been a cornerstone of empirical research on fairness and effectiveness in criminal sentencing. Unwarranted disparities hinder the ability of the criminal justice system to impart equality before the law and undermine its legitimacy (Blumstein, 1982; Kennedy, 1997; Tonry, 1995; Tyler, 2003), which may in turn jeopardize the effectiveness of the criminal justice system and further strain criminal justice relations in the United States (Ruth & Reitz, 2003; Tyler, 1990; Western, 2006). In light of these concerns, it is hardly surprising that sentencing disparities receive so much attention in the empirical literature. Sentencing scholars routinely highlight the importance of conducting critical research on the various forms of unwarranted disparity (e.g. Baumer, 2013; Zatz, 2000). As a result, the fields of criminology, sociology, and law have jointly amassed thousands of studies examining disparities in sentencing outcomes based on defendants' social characteristics, arriving at a general consensus that race, ethnicity, gender, and socioeconomic status all have detectable, though inconsistent, effects on punishment (Baumer, 2013; Daly & Bordt, 1995; Mitchell, 2005; Ulmer, 2012; Wu & Spohn, 2009; Zatz, 2000).

With the introduction of numerous sentencing reforms beginning in the 1980's, it has also become critical to consider the effect of different sentencing schemes on social inequality in sentencing. Zatz (2000) noted the tendency for disparity research in the 1980's and 1990's to focus on determinate sentencing, sentencing guidelines, and mandatory sentencing systems and found that these structured sentencing schemes do not eliminate disparity and in some cases even lead to stronger indirect and interactive effects between social demographics and other factors such as employment status and bail status (e.g. Miethe & Moore, 1985; Spohn et al., 1981). Similar research conducted since 2000 arrives at parallel conclusions (Bushway & Piehl, 2007; Mustard, 2001; Ulmer et al., 2007). Ample work suggests that disparities exist under both structured and non-structured sentencing schemes (Baumer, 2013; Mitchell, 2005; Ulmer, 2012), yet little is known about how more recent developments, such as risk assessments tools and broader evidence-based sentencing practices, affect patterns of inequality in punishment. In line with the work comparing disparities before and after the implementation of sentencing guidelines, determinate sentencing, and mandatory sentencing systems, it is imperative that scholars examine the impacts of evidence-based sentencing on social inequality.

Evidence-based sentencing uses actuarial assessments to help inform sentencing decisions. Actuarial assessment has been defined as the application of "an objective, mechanistic, reproducible combination of predictive factors, selected and validated through empirical research," and applied to key "outcomes that have also been quantified" (Heilbrun, 2009: 133). Actuarial risk assessments in the criminal justice system have been variously used to identify low-risk offenders, good candidates for particular programs, and those at high risk of future violent offending (Cullen & Gendreau, 2000; James, 2015; Latessa & Lovins, 2006; Elek et al., 2015). They have been primarily applied to criminal justice decisions that fall outside the purview of sentencing decisions (e.g. parole decisions), though as many as 20 states have begun to integrate risk assessments into sentencing (James, 2015; Starr, 2014b). These assessments contain a variety of risk, protective, and needs factors, both static and dynamic, that have been shown to influence offender risk and future recidivism (Andrews & Bonta, 2000; Oleson, 2011). The Level of Service Inventory- Revised (LSI-R) is the most popular prediction tool in use among states that have not adopted their own state-specific instruments (Andrews & Bonta, 2000; Elek et al., 2015; Olver et al., 2014).

Criminal sentencing involves a broad range of punishment goals that include retribution, deterrence, incapacitation, rehabilitation, community protection and restoration (von Hirsch, 1976), and the primary aim of evidence-based sentencing is to assist judges in delivering the most appropriate sentences to the most suitable offenders as they address some of these goals (Heilbrun, 2009; Monahan & Skeem, 2015). In particular, risk assessments can be useful for effectively identifying a) which offenders can be given non-custodial sentences without compromising public safety and b) which offenders are at the highest risk of future recidivism (Bergstrom et al., 2009; Hyatt et al., 2011; Warren, 2007). Needs assessments can also be useful for identifying offenders who are well-suited to rehabilitation or restorative justice programming (Andrews & Bonta, 2000, 2002; Monahan & Skeem, 2015; Warren, 2008). This is important because the effective use of incarceration, community punishments and alternative sanctions helps to maximize public resources.

The available empirical evidence suggests that actuarial assessments can serve as valuable predictive tools for judges during the sentencing process (Andrews et al., 2006; Latessa & Lovins, 2010; Skeem & Monahan, 2011; Pew Center on the States, 2011; Casey et al., 2014). The body of research evaluating the predictive validity of risk assessments is often interpreted as evidence that predictions based on risk assessments outperform clinical judgments (Andrews et al., 2006; Gottfredson, 1961; Gottfredson & Gottfredson, 1985; Sacks, 1977), and evidence-based sentencing has been touted as a potential method for alleviating mass incarceration in the United States (Marcus, 2009;

Oleson, 2011; Warren, 2007). The practice may also appeal to citizens because it fosters transparency, emphasizes both objectivity and public safety, and is scientifically validated prior to implementation (Hyatt et al., 2011; Marcus, 2009; Oleson, 2011; Van Nostrand & Lowenkamp, 2013; Warren, 2008; Wolff, 2008).

However, as Stuart and Sykora (2010: 465) opine in their discussion of the flaws in evidence-based sentencing, "when forming public policy to respond to crime, there are no silver bullets. It is an ever-evolving science." Opponents have levelled a variety of criticisms at evidence-based sentencing. Some critics have expressed concerns about the constitutionality of risk assessments, branding them discriminatory and in violation of the Equal Protection Clause (e.g. Starr, 2014b). Others question the ethics of "statistical profiling", equating the use of risk assessments in sentencing to punishment for having similar characteristics to other offenders (Auerhahn, 1999; Hannah-Moffat, 2005). Still others argue that risk estimation is too imprecise to be used with confidence; large margins of error and the possibility of high rates of false positives and false negatives render risk assessment instruments inappropriate for determining individuals' risk of recidivism (Berk et al., 2009; Berk & Bleich, 2014; Cooke & Michie, 2010; Harcourt, 2007; Hart et al., 2007). Finally, many scholars contend that risk assessments in sentencing will exacerbate social inequalities (Hannah-Moffat, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b,c).

This final objection, that risk assessments may yield sentencing decisions that intensify social inequality, echoes repeatedly throughout the evidence-based sentencing literature (Etienne, 2009; Hannah-Moffat, 2005, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b,c) and within the realm of criminal justice policy (e.g. Holder, 2014).

Former Attorney General Eric Holder recently recognized the importance of this issue, noting that "evidence-based strategies show promise in allowing us to more effectively reduce recidivism," but they also hold the potential to "inadvertently undermine our efforts to ensure individualized and equal justice" and to "exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system" (Holder, 2014). The importance of assessing the impact of evidence-based sentencing on inequality cannot be overstated. As Starr (2014c: para. 13) argues, "Criminal justice policy should be informed by data, but we should never allow the sterile language of science to obscure questions of justice." Social equality is a fundamental objective in the pursuit of justice. Thus, the effects on social inequality should be a key concern in the decision to adopt an evidence-based sentencing scheme.

However, the effects of evidence-based sentencing on social inequality remain largely untested in the empirical literature (Heilbrun, 2009). As some policymakers and scholars argue, it is indeed possible that risk assessments are discriminatory toward some groups, such that an evidence-based approach to judicial decision-making will exacerbate sentencing disparities (Holder, 2014; Starr, 2014b,c; Hannah-Moffat, 2005). It is also possible that such an approach will have no discernible effect on disparities or will even reduce them (Laskorunsky, 2017; Monahan & Skeem, 2015). Surprisingly little research exists on the topic. An evaluation that assesses the effects of evidence-based sentencing on social inequality is needed. What is also needed is a more nuanced approach to characterizing social inequality, one which considers the mechanisms through which the use of risk assessments may alter inequality, to provide further elucidation about how evidence-based sentencing alters sentencing outcomes for different groups of offenders.

This dissertation project speaks to this concern by comparing existing patterns of social inequality in Connecticut's sentencing system to patterns of disparity in a simulated evidence-based sentencing scenario. Using data from the Connecticut Judicial Branch Court Support Services Division, it investigates racial, ethnic, gender, and socioeconomic disparity in sentencing among felony offenders convicted in the state of Connecticut between Fiscal Years 2008 and 2010. Then it examines these same disparities in domain-specific and composite scores from the LSI-R risk assessment instrument, as well as in a model that simulates an evidence-based sentencing scheme predicated on legal/case processing factors and offenders' scores from the LSI-R. Finally, it compares the amount of disparity under the current guidelines system to the disparity observed in the simulated evidence-based scenario.

Connecticut provides an interesting context for this evaluation of social inequality and the use of actuarial risk assessments. Connecticut does not have sentencing guidelines, which sets it apart from most of the other states in which studies of sentencing inequalities are conducted. Most evaluations take place in states that have sentencing guidelines (Griffin & Wooldredge, 2006; King & Johnson, 2016; Steffensmeier et al., 1998). Moreover, while the state is politically liberal, its criminal justice system was heavily influenced by the widespread "tough on crime" movement during the 1980's and 1990's and is now dealing with a variety of issues stemming from policies that led to mass incarceration. Connecticut's corrections expenditures have skyrocketed, racial disparities in incarceration rates have grown larger, and the state has felt pressure to reform its court system in order to stem the steady flow of money and offenders into its prisons (Connecticut Sentencing Task Force, 2009). In efforts to make case processing and prison sentences more effective, Connecticut has already incorporated risk assessments into various criminal justice decision-making points, including pretrial detention and bail setting, probation, and parole release decisions. This demonstrates the state's understanding of risk assessments and its openness to evidence-based sentencing as a method for curtailing incarceration, which in turn maximizes the utility of this project for policy decision-making.

The project makes several significant contributions. First it provides a thorough treatment of sentencing disparities in Connecticut, a state with broad offense classifications and no sentencing guidelines. Typical evaluations of sentencing disparity lack measures often thought to be important for judicial decision-making, such as socioeconomic status, family situations, mental health, and attitudes toward crime, but this dissertation capitalizes on measures included in LSI-R risk assessments to capture these and other potentially impactful offender characteristics in its assessment. Second, the project evaluates whether or not the incorporation of risk assessments into judicial decision-making might exacerbate social inequalities, a question frequently asked but never empirically addressed in the criminal justice literature. Third, it explores the mechanisms through which disparities could operate in an evidence-based sentencing scheme, providing information about which domains captured in the LSI-R risk assessment tool are likely to play a role in constraining or amplifying social disparities in sentencing. This information may be meaningful to academics who theorize about the role of various offender characteristics in fostering or counteracting social inequality in punishment, and to policymakers/practitioners who decide when and how to use risk assessments in sentencing.

The proposal proceeds as follows. Chapter 2 considers extant research on racial, ethnic, gender, and socioeconomic disparities in the criminal justice system. Chapter 3 describes the use of actuarial risk prediction in sentencing, including specific mention of the Level of Service Inventory-Revised risk assessment instrument that will be used in this study. Chapter 4 discusses the state of Connecticut and Connecticut's sentencing system as the research context for the study. Chapter 5 addresses relevant theoretical frameworks for evaluating social disparities in sentencing. Chapter 6 details the data and analytic methods that are used in this project. Chapter 7 presents data compilation procedures and descriptive statistics for the project sample. Chapter 8 presents results from all quantitative analyses. And lastly, Chapter 9 closes with a discussion of the results and consideration of study limitations, future directions, and policy implications for this research.

CHAPTER 2: SOCIAL INEQUALITY IN SENTENCING

Unwarranted disparities threaten the legitimacy and effectiveness of the criminal justice system by stimulating perceptions of unfairness and promoting public distrust (Tonry, 1995; Western, 2006), which is why social inequality is perhaps the most widely-studied point of inquiry in empirical sentencing research (e.g. Baumer, 2013; Bontrager et al., 2013; Bushway & Piehl, 2001; Mustard, 2001; Ulmer, 2012; Zatz, 2000). In order to situate the investigation into Connecticut sentencing within the context of prior evidence on social inequality in sentencing, this chapter provides an overview of the extant research on sentencing disparities with a specific focus on the effects of race and ethnicity, gender, and socioeconomic status. It additionally reviews some of the limitations that characterize this body of research.

Racial and Ethnic Disparities

The body of research on racial disparities in sentencing is voluminous, and the evidence suggesting a disadvantage for minority offenders at sentencing is compelling (Baumer, 2013; Mitchell, 2005; Pratt, 1998; Spohn, 2000; Ulmer, 2012; Zatz, 2000). In one review of the literature, Baumer (2013: 15) proclaims that “There is little disagreement among scholars about whether there is racial inequality in sentencing outcomes… there are considerable racial inequalities that are highly visible in the final sentencing stages, and these inequalities are quite large.” Notably, however, the effects of race on sentencing outcomes are not universal; though relevant research collectively yields an affirmative answer to the question of whether or not disparity exists, equally relevant questions include when, where, and how these disparities can be observed. These complementary queries have been the subject of much research as well. As Baumer

(2013) and others also note, the meaning of these inequalities are less definitive. Much like the role of race in American society, the role of race at sentencing is highly nuanced. The complexity of sentence decision-making combined with the hypothesized multifunctionality of race as both a demographic characteristic and an indicator of social and economic status, behavioral tendencies, and even a legacy of historical oppression muddies interpretations of the relationship (Howard, 1975; Steffensmeier et al., 1998).

Several reviews provide a good synopsis of the race and sentencing literature (Baumer, 2013; Mitchell, 2005; Pratt, 1998; Spohn, 2000; Ulmer, 2012; Zatz, 1987; 2000). The earliest studies were documented by Zatz (1987), who traced the progress of sentencing research through four historical waves with increasingly more rigorous methodologies and better acknowledgement of the subtleties that characterize the race and sentencing relationship. Spohn (2000) reviewed 40 studies that investigated the link between race/ethnicity and sentence severity, concluding that while there is a direct link between race and punishment decisions, contextual and indirect effects are also key parts of the equation. Interestingly, Pratt's (1998) meta-analysis of studies examining the effect of race on sentence length found no significant disadvantage for black defendants. That said, the analysis did not consider effects on the incarceration decision, and race effects tend to be stronger for that decision than for sentence length (see Mitchell, 2005).

In a more comprehensive meta-analysis of 116 black-white contrasts drawn from 71 empirical studies, Mitchell (2005) identified a small but statistically significant direct effect of race on sentencing, such that African-American defendants received harsher punishment than white defendants. 73% of the effect sizes identified in the study showed a black disadvantage. Mitchell's (2005) evaluation also went one step further, examining the contexts in which studies were more likely to yield a significant effect size. He found that racial disparities were largest among non-federal drug offenses, in imprisonment and discretionary punitiveness (e.g. upward departures, sentencing enhancements) decisions, in jurisdictions with structured sentencing regimes, in cases collected from a single city or county, in studies with fewer control variables and less precise measures of criminal history and offense seriousness, and in studies that were published. Given the substantial variability in these characteristics across sentencing studies, Mitchell's (2005) conclusions are of paramount importance. A full understanding of racial disparities in sentencing requires thoughtful consideration of the settings in which scholars are able to observe them.

More recently, Baumer (2013) reassessed the sentencing literature, observing that research tended to focus on one or more of four primary objectives: detecting racial disparities, detecting racial discrimination, evaluating the effects of policy interventions on racial disparities, or assessing the impact of race on legal decision-makers. In concert with the reviewers before him, he also noted the importance of race's conditional role in sentencing and identified several shortcomings, both conceptual and methodological, that have limited the growth and development of the research.

Even so, research on race and sentencing disparities has made advancements down a variety of topical avenues in recent years. As Johnson and Lee (2013) note, the incorporation of new dependent variables is particularly valuable in the context of modern sentencing systems in which judicial discretion is no longer the sole determinant of punishment outcomes. Sentencing guidelines, mandatory minimums, habitual offender laws, and other contemporary sentencing reforms have altered the structure of sentencing

(Tonry, 1996), such that punishment may be thought of as the confluence of several decisions, ranging from guidelines placement to the imposition of enhancement statutes, made by sentencing judges, prosecutors, and even policymakers who craft sentencing policy (Bushway & Forst, 2013). Failure to acknowledge the role of related courtroom decisions in determining punishment ignores the mechanisms through which sentencing is determined. For example, Mustard (2001) found that black defendants received substantially longer sentences than white defendants, but also that over half of the disparity was due to guidelines departures. Thus, this new line of research is exemplified by work on disparities in mandatory minimums (Crawford et al., 1998; Ulmer et al., 2007), guidelines departures (Britt, 2000; Johnson, 2005), and alternative sanctions (Gainey et al., 2005; Johnson & DiPietro, 2012), much of which bolsters the notion that racial and ethnic disparities are found in a variety of discretionary decisions that in turn impact sentencing outcomes. Evaluating disparities in evidence-based sentencing, which allows actuarial risk assessment to influence risk perceptions and ultimately punishment outcomes, fits in well with this direction of development.

Expanding the conceptualization of race to consider more than just black-white dichotomies has also resulted in more fine-tuned conclusions about the role of race in sentencing. Studies examining the effects of skin tone and the Afrocentricity of facial features find that darker skin tone and Afrocentric features contribute to harsher punishment above and beyond the influence of categorical race (Burch, 2015; Eberhardt et al., 2006; King & Johnson, 2016). There is also a growing body of evidence that racial and ethnic groups other than blacks experience advantages or disadvantages at sentencing. Albonetti (2002), Doerner and Demuth (2010), Holmes and colleagues

(1996), Spohn and Holleran (2000), and Steffensmeier and Demuth (2001) all highlight the importance of evaluating disparities for Hispanic defendants, while others point to a need for more investigation into sentencing for other racial groups such as Asians and Native Americans (Alvarez & Bachman, 1996; Franklin, 2013; Johnson & Betsinger, 2009). On top of these effects, some studies find that immigration/citizenship status exerts an additional influence on sentencing (Demuth, 2001; Hartley & Tillyer, 2012; Wolfe et al., 2011). This body of work demonstrates that moving beyond simple black-white comparisons, as this project will do, holds great potential for further clarifying racial/ethnic and related disparities in sentencing.

Gender Disparities

Relative to racial and ethnic disparities, gender disparities in sentencing have received less scrutiny in the empirical literature. Historically, women have constituted only a small fraction of known offenders (Bonczar, 2003) and evaluations of their treatment in the criminal justice system have been viewed as a niche enterprise falling under the umbrella of "feminist criminology" (Chesney-Lind, 1997). Early reviews of the gender disparities literature observed this inattention (Bickle & Peterson, 1991; Daly & Bordt, 1995; Nagel & Hagan, 1983). Nagel & Hagan (1983) identified only 16 studies in which gender was used as a predictor of sentencing outcomes, while Daly and Bordt (1995) included only 28 sentencing studies in which gender was the primary focus. Both reviews found evidence that women tend to receive lighter punishments at sentencing than men (Nagel & Hagan, 1983; Daly & Bordt, 1995). While effect sizes were relatively small when compared to the effects of legal factors such as criminal history, they were quite consistent. Some of the research detected no gender disparities, but the majority identified a disparity that advantaged women at the sentencing stage. These findings went largely uncontested, as they fit with the conventional view of women as less dangerous and less responsible for their actions, less suitable for prison due to family responsibilities, and better targets for rehabilitation efforts (Bickle & Peterson, 1991; Daly, 1989). As Crew (1991) opined, gender disparities in sentencing were treated as commonplace and legitimate.

More recently, the rapid growth of female representation in the criminal justice system (US Department of Justice and Federal Bureau of Investigation, 2009) has seemingly begun the process of integrating gender differences into the mainstream criminal justice conversation. A meta-analysis of contemporary research on gender and sentencing identified 58 relevant studies published between 1991 and 2011, 39 of which were published in the 2000's (Bontrager et al., 2013). The review found clear evidence that women remain an advantaged population at sentencing in modern times, especially for the decision to incarcerate, but with one important caveat: studies that used data collected in or after 2000 showed smaller gender gaps in sentencing than studies that used data collected in the 1990's. The authors note that this shrinking is consistent across sentencing outcomes, going so far as to conclude that "women no longer enjoy significantly shorter sentences, have lower odds of incarceration, or have better chances at a sentencing departure than their male counterparts" (Bontrager et al., 2013: 366). Still, generally speaking, contemporary research on gender disparities continues to identify a small advantage for women at sentencing while both improving methodologically and becoming more attuned to the nuances of inequalities in sentencing. Evidence indicates that this is true in state courts (Griffin & Wooldredge, 2005; Steffensmeier et al., 1993) as well as in federal courts (Albonetti, 2002; Brennan & Spohn, 2009; Doerner & Demuth, 2010; Mustard, 2001), that the effect seems to be larger in the decision to incarcerate than in the determination of sentence length (Bontrager et al., 2013), and that there are also several interactive effects that color the relationship between gender and punishment (Bloch et al., 2014; Doerner & Demuth, 2010).

Sentencing scholars have made great strides by acknowledging that gender does not operate in isolation but rather intersects with other conditions to create advantages or disadvantages for certain defendants at sentencing. A noteworthy portion of the literature on gender disparities in sentencing has thus been devoted to potential interactive effects between gender and other defendant or case characteristics. Interactions between gender, race, age, and class are discussed later in this chapter. Studies examining the joint effects of gender and familial factors, factors such as marital status and having children or other dependents, is somewhat mixed. Koons-Witt (2002) found that female offenders were granted leniency if they had dependent children, while others find that neither marital status nor responsibility for dependents condition the effects of gender (Brennan & Spohn, 2009; Griffin & Wooldredge, 2006; Mustard, 2001) or even that the benefits from being a family caretaker are larger for male than for female defendants (Freiburger, 2010).

A second area of study is the conditional effects of offense type on gender disparity. It is well-documented that women tend to commit fewer and less serious crimes (Belknap, 2007; Britton et al., 2017), and several scholars have theorized that women who commit more serious or violent crimes are punished similarly to men because they have violated both law and gender norms (e.g. Franklin & Fearn, 2008). Moreover, others postulate that high offense severity may constrain judicial discretion by requiring harsh sentences, such that there is more gender uniformity in sentencing among serious offenses (see Kalven & Zeisel, 1966 for the original formulation of the liberation hypothesis). In line with these expectations, several studies find that females receive leniency relative to males among those committing less serious crimes (Koons-Witt et al., 2014; Rodriguez et al., 2006); however, others identify little to no discrepancy across offense types (Mustard, 2001). Much like interactions between race and gender as well as race and familial factors, the nature and import of interactions between gender and offense type therefore remain in question.

Socioeconomic Disparities

The importance of social class for determining sentencing outcomes has been a hallmark assumption of criminal justice literature for decades. Conflict models of criminal justice posit that the law is applied discriminately as a means of oppressing marginalized groups, namely blacks and the lower class, and preserving the power of the social elite (Chambliss & Seidman, 1971; Quinney, 1973). Criminal justice sanctions are one of the mechanisms that function as a tool of institutional oppression, and as a result, lower class criminal defendants receive harsher punishments than their upper class counterparts (Chambliss & Seidman, 1971). This proposition, that the lower class is disadvantaged at sentencing, is widely acknowledged (D'Alessio & Stolzenberg, 1993; Chiricos & Waldo, 1975; Kramer & Ulmer, 2009; Miethe & Moore, 1985). However, the impact of social class on sentencing outcomes remains largely an assumption. Countless works reference the potential importance of SES as an independent or mediating influence and hail it as a crucial topic for future sentencing research (e.g. Ulmer, 2012; Zatz, 2000), but it predominantly plays a supporting role in the empirical sentencing literature.

This is likely because measures of social class are often coarse, incomplete, or nonexistent. SES is widely accepted as "a composite measure that typically incorporates economic status, measured by income; social status, measured by education; and work status, measured by occupation" (Dutton & Levine, 1989: 30). Rarely are all three of these components- income, education, and occupation- captured in official court records, making it difficult for scholars to fully account for SES in sentencing research. Studies that do include robust measures of SES are few (for exceptions, see D'Alessio & Stolzenberg, 1993; Chiricos & Waldo, 1975). Instead, the modal method is to include one or another of the SES components as the only indicator. In other instances, researchers use other variables hypothesized to relate to social status as proxies for SES, such as defendants' type of attorney (Bloch et al., 2014; King & Johnson, 2016; Kutateladze et al., 2014), or do not include any socioeconomic indicators at all (Britt, 2000; Bushway & Piehl, 2001).

Educational attainment is commonly used as a stand-alone indicator of SES. While defendants' education may indeed serve as an indicator of social status, judges may also use educational background as a signal for defendants' work ethic, commitment to prosocial achievement, and desirability as a job applicant, as qualitative work in the Pennsylvania court system suggests (Kramer & Ulmer, 2009). Franklin's (2017) sentencing evaluation put the spotlight squarely on educational attainment, finding that having a high school degree significantly impacted both the in/out and sentence length decisions and even moderated the effects of other extralegal variables such as race and gender in some cases. Obtaining a college degree did not have any significant effect. Franklin's (2017) findings are consistent with much of the work on education and sentencing. There appears to be a general consensus that receiving a high school diploma provides an advantage at both sentencing stages (Albonetti, 1997; Franklin, 2013; Johnson & Betsinger, 2009; Mustard, 2001; Rehavi & Starr, 2014), while the evidence for significant effects of going to college or obtaining a college degree is on the whole much weaker (Franklin, 2013; Johnson & Betsinger, 2009; Koons-Witt, 2002; & Starr, 2014).

Employment factors are also often used in isolation to represent SES. Early sentencing studies incorporated indexes measuring occupational prestige (Bedau, 1964; Wheeler et al., 1982; Willick et al., 1975) and often obtained little evidence that defendants' occupation significantly impacted sentencing. On the other hand, there is some evidence that unemployed offenders receive harsher sentences than employed offenders (Chiricos & Bales, 1991; Myers, 1987; Wooldredge, 2010). However, it is worth noting that the effects of employment status also appear to vary across contexts, including geographical contexts (Nobiling et al., 1998), sentencing guidelines models (Koons-Witt, 2002; Miethe & Moore, 1985), judges (Anderson & Spohn, 2010; Wooldredge, 2010), and offender demographic characteristics (Chiricos & Bales, 1991; Spohn & Holleran, 2000).

Income, the third traditional component of SES, is more difficult to obtain in criminal court datasets and therefore a less common control in the empirical literature than education or employment. To illustrate, Hagan and colleagues' (1980) analysis of white collar sentencing included a dichotomous measure indicating whether each defendant had an income of over or under $13,777 per year, but the authors opted to present results focusing on divisions of educational attainment instead. Mustard's (2001) evaluation of sentencing outcomes for federal offenders included a series of dummy variables capturing defendants' income ranges, ultimately concluding that defendants who earned less than $5,000 per year were the most disadvantaged at sentencing. Wooldredge (2010) took a different approach to accounting for income by measuring whether or not defendants were reliant on relatives, friends, or the government for financial support and found that such reliance was related to higher chances of being incarcerated. Though these studies and others together lend some credence toward the idea that earnings are negatively associated with punishment severity, the dearth of research on the topic precludes firm conclusions.

Finally, a few other case characteristics have been used as proxies for SES in sentencing evaluations. Type of defense attorney is occasionally used as a rough indicator of defendants' income and wealth; representation by a public defender or court-appointed attorney is presumed to signify an indigent defendant, while representation by private counsel requires at least some financial means (Skolnick, 1967). Sentencing studies that use type of attorney as a proxy for SES tend to find small but statistically significant effects of attorney type on sentencing outcomes (King & Johnson, 2016; Kutateladze et al., 2014; Nagel, 1969). Others have used income characteristics of the neighborhoods in which defendants were arrested (Kutateladze et al., 2014) or counties in which defendants live (Rehavi & Starr, 2014) as additional proxies for SES, likewise concluding that these variables have slight but significant effects on sentencing.

Intersectionality in Sentencing Disparities

In addition to observing a wide range of direct and independent effects of social demographics on sentencing, scholars have begun to recognize the need to examine interactive effects. The notion of intersectionality originated in black feminist scholarship, where Crenshaw (1989; 1991) and others argued that people's identities are multidimensional, and each facet of a person's identity helps construct their unique experiences and shape their treatment by others (Davis, 1981). For example, race and gender interact, such that black women have different experiences both from white women and from black men. In the context of the criminal justice system, this suggests that not all defendants of the same race/ethnicity, gender, or class have the same experiences and receive the same advantages or disadvantages during case processing, so it is important that sentencing research addresses the intersection of various racial/ethnic, gender, and class inequalities. Though the body of research focusing on intersectional disparities in formal punishment is still relatively sparse, it is steadily growing, and results generally indicate that the combined effects of various group memberships are impactful above and beyond the effects of each membership individually.

To illustrate, a noteworthy portion of the literature on joint disparities in sentencing has been devoted to potential interactive effects between race and other defendant or case characteristics. One especially prevalent focus of study has been the intersection of gender and race, or what contemporary feminist scholars refer to as multiracial feminism (Burgess-Proctor, 2006). Research in this area indicates somewhat equivocally that gender and race may both exert contextual influences on each other, with the largest penalties given to minority men (Albonetti, 2002; Brennan & Spohn, 2009;

Doerner & Demuth, 2010; LaFrentz & Spohn, 2006; Spohn & Holleran, 2000; Steffensmeier et al., 1998; Steffensmeier et al., 2016), though some work instead finds that the gender gap is comparable across racial and ethnic groups (Bloch et al., 2014; Spohn & Beichner, 2000) or that racial gaps are comparable between male and female offenders (Steffensmeier et al., 1998).

Building on this literature, the notion that age interacts with race and gender to procure harsh punishments for particular offenders has garnered considerable empirical support. Scholars repeatedly identify young black males, and occasionally young Hispanic males as well, as a group for whom age is especially impactful and for whom punishment is often the most severe (Doerner & Demuth, 2010; Nowacki, 2016; Steffensmeier et al., 2016; Spohn & Holleran, 2000). Steffensmeier and colleagues' (2016) assessment of intersectional inequality in Pennsylvania produced solid evidence that both young black males and young Hispanic males received the harshest sentences, while young females and some of the oldest defendants received leniency. It is often posited that widespread stereotypes portraying young minority men as dangerous influence judicial decision-making and result in harsher punishments for young, minority, male offenders than any other offending group (Spohn & Holleran, 2000; Steffensmeier et al., 1998).

The intersection of class with race/ethnicity and gender is infrequently explored in the sentencing literature, though select evaluations show some contextual differences among these social characteristics. Spohn and Holleran (2000) found that unemployed black and Hispanic males were more likely to be imprisoned than employed white males, but the relationship did not hold for female defendants. LaFrentz and Spohn's (2006) evaluation of federal drug sentencing identified effects of employment status that were conditioned by race/ethnicity; employment benefited white defendants at sentencing but had no effect for blacks and Hispanics. Mustard's (2001) study of the federal system found evidence that the white-Hispanic disparity was conditioned by education; the ethnic disparity was smaller among more educated defendants. Additional work on the joint influences of class, race, gender, and age may eventually explain more of these complex relationships.

Methodological Limitations of Sentencing Disparity Research

Overall, the research on disparities in sentencing has advanced greatly in recent decades, clarifying the extent of sociodemographic advantages and stimulating new and important inquiries into the treatment of various types of offenders in the criminal justice system. But sentencing research still has several methodological challenges. Some of these difficulties are specific to research on a single sociodemographic factor, while others characterize most or all of the research on race/ethnicity, gender, and SES disparities. The rest of this discussion emphasizes the latter type and focuses on how this dissertation project seeks to overcome some of these challenges.

First, concerns about model misspecification due to omitted or poorly-measured variables permeate the sentencing literature. For example, among studies focusing on racial disparities, robust measures of SES, which is closely intertwined with race, are few and far between. As a result, it remains unclear whether observed racial disparities are truly due to racial group membership or whether they are in fact the result of biases against the lower class (Baumer, 2013; Mitchell, 2005). Because minority defendants also have substantially longer and more serious criminal histories than white defendants

(Kutateladze et al., 2015), it is important to capture detailed information about defendants' criminal histories in sentencing analyses.

Similarly, critics of gender disparity evaluations often point to a lack of fine-grained offense and criminal history information in sentencing research, arguing that sentencing leniency for women may be due to qualitative differences in male and female offending that are not captured in traditional court datasets (Bontrager et al., 2013; Steffensmeier et al., 1995). More specific offense type information helps inform this critique. Many also speculate that gender disparities may be due to "gender-related variables" such as marital status and responsibility for dependents that are often missing from analyses (Daly, 1989; Koons-Witt, 2002). As discussed previously, the biggest concern with studies of social class disparities is that they seldom fully capture defendants' social class (Ulmer, 2012; Zatz, 2000). Incomplete measurement of the factors playing into each defendant's social standing precludes a full and nuanced understanding of how status plays into sentencing outcomes. Additionally, virtually none of the work on race/ethnicity, gender, or SES inequalities account for prosocial or criminal attitudes, demeanor in the courtroom, remorse, or lifestyle characteristics such as residential stability and family life. These factors are not included in official court records but have the potential to significantly impact sentencing and even explain some part of observed social inequalities.

This dissertation project improves upon previous sentencing inequality research first by including a wide variety of sentence predictors and addressing many of the concerns about omitted variable bias when modeling sentencing outcomes. Criminal history measures are numerous, accounting for juvenile and adult misconduct, incarceration, escape histories, institutional misconduct, and prior violence. Measures of employment, education, and financial stability all contribute to a more robust representation of SES, and family/marital statuses help clarify observed gender effects. Furthermore, the inclusion of variables capturing attitudes/orientations toward crime and lifestyle characteristics such as residential situation, leisure activities, criminal companionship, alcohol/drug problems, and mental health allows for much more nuanced assessments of the indirect pathways through which race, ethnicity, gender, and SES influence punishment.

A second concern in the social inequalities literature is the small number of court systems from which sentencing data are drawn. Specifically, the majority of studies on sentencing disparities sample from only a few select states, most of which have sentencing guidelines, or from the federal court system. This tendency constrains the generalizability of findings from the research. Each state court system is unique (for one example of a way in which state courts vary, see Kauder and Ostrom’s (2008) profile of state sentencing guidelines), as is the federal court system, and disparities observed in one state’s courts may not be observable in another’s. This dissertation project focuses on sentencing in Connecticut state courts, which are not governed by any type of sentencing guidelines and have not been subjected to much empirical scrutiny (Gertz & Price, 1985; Zeisel & Diamond, 1977).

The third issue relates to the limited incorporation of social inequality into sentencing policy evaluations, specifically the failure of evaluations to consider a wide range of potential social disparities associated with policy implementation. This point is demonstrated clearly in increasingly popular “evidence-based approaches” to policymaking. Evidence-based policymaking is based on the principle that decision-making is better if it is informed by rigorous research. Policymakers put a premium on information that tells them how effective particular policies are and helps them determine how available resources can be used to maximize utility and do the most good (Head, 2010; Pew-MacArthur Results First Initiative, 2014). In practice, the evidence policymakers seek first and foremost is that which informs a policy's return on investment (e.g. Drake et al., 2009). To illustrate, the Justice Reinvestment Initiative, run by the US Office of Justice Programs' Bureau of Justice Assistance, maintains a primary goal of using evidence-based decision-making to improve public safety through the allocation of resources in cost-beneficial ways (La Vigne et al., 2014). Trapped by perennial budget constraints, policymakers want to identify and ultimately implement policies that are the most cost-effective. Financial costs, however, are not the only type of costs that a specific policy may incur.

Policy evaluations do not typically factor in a sufficiently wide range of social costs when determining costs and benefits, likely because of the difficulties associated with collecting or obtaining the data necessary to evaluate these social costs. A few states, including Connecticut, have passed legislation mandating the consideration of racial impact statements, which assess the racial impact of proposed changes to sentencing and parole policies (London, 2011; Mauer, 2008). Aside from these racial impact statements, though, the potential for disproportionate impact on marginalized groups is seldom seriously considered. When it is, it is often as an afterthought (for exceptions, see Crawford et al., 1998; Miethe & Moore, 1985; Ulmer et al., 2007). The research on evidence-based sentencing is no exception. The body of research that evaluates evidence-based sentencing focuses primarily on the overall predictive validity of risk assessment instruments and the impact that evidence-based sentencing will have on correctional populations and budgets (Hyatt et al., 2011; Ostrom et al., 2002; Ruback et al., 2016). The potential for social inequality to worsen as the result of evidence-based sentencing has not been assessed empirically (Laskorunsky, 2017). This dissertation project first and foremost targets this question, serving as an impact analysis for race, ethnicity, gender, and SES effects prior to the implementation of evidence-based sentencing in Connecticut.

In sum, research on social inequality in punishment is prolific. The literature examining racial, ethnic, gender, and SES disparities has found that these sociodemographic factors exert both direct and indirect effects on sentencing. While legal factors such as criminal history and offense severity account for the majority of the variation in sentencing outcomes, the presence of these unwarranted disparities still raises concerns about fairness in the criminal justice system. However, research on punishment inequality is subject to a variety of limitations that leave ample room for improvement. This dissertation addresses these challenges by assessing both current social inequality and the potential effects of evidence-based sentencing on inequality in Connecticut. The next chapter provides an overview of evidence-based sentencing.

CHAPTER 3: EVIDENCE-BASED SENTENCING

With over 1% of the American population housed in jail or prison (Carson & Anderson, 2016), the United States' culture of mass incarceration has become unsustainable (Cullen et al., 2011), and excessive prison populations have incited demands for new sentencing policies that curb incarceration rates without jeopardizing historically low US crime rates (Warren, 2007; 2008). From both sides of the political aisle, evidence-based sentencing has been touted as a promising response to these demands (Arnold & Arnold, 2015; Marcus, 2009). This chapter describes risk assessments and evidence-based sentencing, with special consideration of the risk assessment that is used in this study's evaluation of evidence-based sentencing, the Level of Service Inventory-Revised (LSI-R).

Risk Assessments

Evidence-based sentencing is sentencing that relies on the use of actuarial assessments (Heilbrun, 2009). Evidence-based sentencing instruments are typically based on previous regression analyses linking recidivism rates with various offender and offense characteristics; each individual offender is assessed on variables that have been empirically associated with recidivism, such as criminal history and SES, and then item scores are summed to produce a score that serves as an indicator of overall reoffending risk. The foundation of evidence-based sentencing is the targeting of risk; risk assessments are intended to identify offenders' risk of reoffending, and judges are expected to use those evaluations of risk to assign proper sentences.

The notion of offending risk is hardly foreign to the American criminal justice system (Harcourt, 2007). As Monahan and Skeem (2014) note, there is a long tradition of risk prediction in the criminal justice system, dating back to the late nineteenth century when states began adopting community supervision programs, like probation and parole, that were both assigned and terminated based on assessments of offenders' risk of recidivism. What sets risk assessments used in contemporary evidence-based sentencing practices apart from past risk prediction techniques, advocates argue, is the incorporation of empirically sound, objective information about recidivism and interventions into risk calculations (Etienne, 2009; Marcus, 2009; Oleson, 2011; Warren, 2007). Rather than relying on potentially capricious intuitions and unsystematic judicial considerations of offender and offense characteristics to dictate sentences, evidence-based sentencing is intended to rely on social scientific evidence that can be used to maximize public safety.

Bonta (1996) illustrates this distinction by characterizing three "generations" of risk assessments in the criminal justice system. First-generation risk assessments are quasi-clinical interviews in which criminal justice officials use offenders' answers to unstructured questions to determine risk. These types of assessments, based solely on professional judgment, are still used in some jurisdictions today, but they have been replaced largely by second- and third-generation assessments. Second-generation risk assessments generally retain the interview format but improve upon their predecessors by incorporating actuarial instruments into decision-making. The corrections official is thus informed about offender characteristics that had been empirically shown to relate to risk of recidivism. However, second-generation assessments consist primarily of static risk factors that cannot be eliminated (e.g. age, criminal history, gender). While good predictors of recidivism, static risk factors only aid in managing offenders' risk; they provide little guidance to criminal justice actors about how to *reduce* risk. This is the problem that third-generation risk assessments attempt to tackle. Third-generation risk assessments, also commonly referred to as risk-needs instruments, consider both static and dynamic risk factors in their evaluation. These assessments, which encompass several of the most popular risk prediction instruments in use, acknowledge that offenders' circumstances and needs can change over time and that individual offenders' risk can be reduced if the proper dynamic risk factors are addressed (Andrews & Bonta, 2000; Oleson, 2011). Since Bonta's (1996) original formulation, a fourth generation of risk assessments has been added to the fray. Fourth-generation risk assessments integrate case planning tools into the instrument; in addition to evaluating risk, they acknowledge and recommend supervision and intervention options (Andrews et al., 2006).

Risk assessments were first introduced in the correctional system, as a means of evaluating incarcerated offenders' risk of recidivism and potential for rehabilitation (Andrews & Bonta, 2000; Bonta, 1996), but their use has expanded into criminal sentencing. At least 20 states have now incorporated risk assessment instruments into the sentencing process in some way (Starr, 2014b). Its popularity in both courts and correctional systems may be due in part to the variety of identifications for which risk assessments can be useful. First, these assessments can be used to identify which offenders can be given non-custodial sentences without jeopardizing the safety of the community. For example, the Virginia Criminal Sentencing Commission developed a Risk Assessment Instrument designed to identify the best low-risk offender candidates for community supervision and alternative sanctions programming. The goal was to divert a full 25% of nonviolent offenders who would otherwise have been incarcerated (Kleiman et al., 2007; Ostrom et al., 2002). This tactical use of risk assessments in particular is designed to increase resource optimization by reducing expenditures on those offenders who are least likely to require incapacitation in order to be crime-free.

On the other side of the coin, risk assessments can be used to identify which offenders have a relatively high risk of future recidivism and pose the greatest threat if released (Bergstrom et al., 2009; Hyatt et al., 2011; Warren, 2007). Though evidence-based sentencing programs are thought to be most effective among moderate and high-risk offenders, this purpose is less often an explicit objective of risk assessment tools, as its inherent orientation toward harsh sentencing (even if for only a small proportion of all offenders) seemingly belies the shift in academic and policy rhetoric toward reduced sentences and alternatives to prison. Framing evidence-based sentencing as a means of mitigating sentences and reducing incarceration is often considered a more effective way to argue for it (Etienne, 2009; Starr, 2014b). Among third- and fourth-generation instruments that include dynamic risk and need factors, the assessments can also be used to identify offenders who are well-suited to particular rehabilitative treatments or restorative justice programming (Andrews & Bonta, 2000, 2002; Monahan & Skeem, 2015; Warren, 2008). The development of numerous evidence-based alternative sanctions and community release conditions has pushed practitioners to consider the types of offenders for which each program is most effective, and risk-needs assessments can assist in the identification of suitable offenders.

The distinct generations of risk assessments illustrate one set of ways in which the tools vary, but risk assessments can differ substantially in other ways, even within generational groupings. For instance, risk assessments may be intended to predict different forms of negative behavior. Some risk assessments are designed to predict overall recidivism (Coulson et al., 1996; Skeem & Lowenkamp, 2016), while others may be used specifically to predict more specific types of reoffending such as domestic violence, general violence, property crimes, or drug offenses (Dutton & Kropp, 2000; Yang et al., 2010). Risk assessments also have different target offending populations. Risk assessments have been developed for the general offending population, as well as for more specific groups such as juveniles, females, and sex offenders (Hanson, 1997; Hoge, 2002; Voorhis et al., 2010). Further, the domains included in risk assessment instruments vary both in number and type. The number of individual items/questions on risk assessment tools can range from less than 10 to over 100, and though there are some areas covered by most mainstream risk assessment tools (e.g. criminal history, antisocial behavior, substance abuse), other domains such as demographics, recreational activities, and neighborhood or residential characteristics appear less frequently (Oleson et al., 2011). Other sources of variation may include the information collection methods (e.g. semi-structured interview, official record reviews, self-report surveys), the populations on which tools were initially validated (e.g. convicted male offenders, Canadian offenders, general population), and whether the tools are proprietary or available for public distribution.

Pros of Evidence-Based Sentencing

In spite of the numerous variations in risk assessment tools, the majority of these tools achieve comparable benefits and are subject to many of the same limitations. Generally, the use of risk assessments in sentencing is attractive for several reasons. First, it constitutes a data-driven, scientifically validated approach to sentencing (Etienne, 2009). Utilitarian risk prediction is already a core component of judges' sentencing decisions, so advocates argue that evidence-based sentencing can simply add a layer of precision to sentencing practices without constraining discretion (Kern & Bergstrom, 2013; Oleson, 2011). Most risk assessment instruments contain only items that have been empirically shown to relate to recidivism, and the accuracy with which the instruments predict recidivism is generally tested using samples of convicted offenders before the instrument is fully integrated in sentencing (Ostrom et al., 2002). Thus, evidence-based sentencing offers a more rigorous version of risk prediction than what judges already use to make decisions. Bolstering this argument are reports that risk assessment instruments outperform human judgment, such that predictions made using risk assessments have greater predictive validity than clinical assessments (Andrews et al., 2006; Gottfredson, 1961; Gottfredson & Gottfredson, 1985; Sacks, 1976).

A second advantage of evidence-based sentencing is its potential to reduce incarceration without endangering public safety. As discussed at the beginning of this chapter, evidence-based sentencing is in part a pragmatic response to harsh sentences and mass incarceration in the United States; it is viewed as a method for decreasing reliance on the American prison system by diverting low-risk offenders away from incarceration (Marcus, 2009; Warren, 2007). Additionally, by matching offenders with punishment plans that will reduce their likelihood of reoffending, evidence-based sentencing can increase the resource efficiency of the criminal justice system and maximize its impact on crime. As a result, evidence-based sentencing also enjoys support from a broad swathe of the criminal justice community (Oleson, 2011; Warren, 2008). Sentencing judges in particular believe that evidence-based sentencing is worth pursuing as a mode of increasing the effectiveness of punishment (Peters & Warren, 2008).

Relatedly, evidence-based sentencing is often lauded for its focus on public safety and rehabilitation rather than retribution (Warren, 2007). In a retributive framework, offenders are punished simply because they deserve it. The punishment's primary function is to be the consequence of an offender's behavior (Monahan & Skeem, 2016). This penal philosophy was one of several that contributed to punitive sentencing and the rise of mass incarceration in the 1980's and 1990's (Gertner, 2010). Now that incarceration rates are becoming unsustainable, utilitarian "smart on sentencing" approaches that emphasize crime reduction, public safety, and rehabilitation are taking center stage among academic and practitioner communities (Etienne, 2009; Marcus, 2009; Warren, 2007). Evidence-based sentencing embodies this emerging ideology, representing the incorporation of risk reduction goals and rehabilitation back into sentencing.

Finally, the use of risk assessment instruments introduces a mechanism for increasing objectivity and transparency in sentencing decisions. For decades, court actors have been called out for making decisions behind closed doors and then failing to justify them publicly, a complaint that erodes public trust in the criminal justice system and threatens its legitimacy (Bibas, 2006). The use of innovations that generate systematic, reproducible outputs, such as risk assessments that calculate risk scores, may thus appeal to those who are concerned about personal biases influencing courtroom outcomes. Much like sentencing guidelines that were intended to limit judicial discretion and curtail bias in sentencing, risk assessments can be viewed as an accessible decision-making 'guide' that promotes the use of open, objective offender evaluations.

Cons of Evidence-Based Sentencing

At the same time, there have also been several criticisms levelled at the use of risk assessments. First, despite its potential to increase transparency, evidence-based sentencing often suffers from a lack of both transparency and public understanding. Because many risk assessment instruments were developed by private entities that retained proprietary rights, the factors and algorithms used to calculate risk scores cannot always be shared with the judges who use them and the defendants who receive them. This creates an aura of secrecy around evidence-based sentencing and can make it difficult for defendants, the public, and sometimes even court actors to trust risk assessments. One good example of this phenomenon occurred in Pennsylvania, where a risk assessment instrument was created to assist in the state's parole decisions. In detailing the instrument's development and implementation, Bergstrom and colleagues (2009) describe how the tool was not initially made public, which led to widespread public distrust and ultimately a loss of support for the parole system.

A recent court case in Wisconsin also demonstrates this issue. In *Loomis v. Wisconsin*, Eric Loomis argued that using proprietary risk assessments to make sentencing decisions violates defendants' constitutional right to due process; the proprietary nature of the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), a risk assessment that a Wisconsin judge relied on to issue Loomis' six-year prison sentence, prevented Loomis from challenging the validity and accuracy of the assessment and his sentence (*State v. Loomis*, 2016). Though the Wisconsin Supreme Court ruled that due process rights were not violated in the case (and the US Supreme Court declined to issue a writ of certiorari upon Loomis' appeal), the dispute brought to light concerns about the use of proprietary risk assessments that may well spur other legal challenges in the future.

Second, some argue that risk assessments are too imprecise in their estimations of recidivism risk to warrant their use at sentencing. Like other predictive tools, recidivism risk assessments are not always accurate; empirical tests indicate that they have large margins of error, such that some offenders labeled “high-risk” by risk assessment tools do not recidivate while some offenders labeled “low-risk” do (Fass et al., 2008; Gendreau et al., 2002; Wong & Gordon, 2006). Inaccurately assessed defendants will either suffer needlessly or pose a significant risk to the public, so precision is paramount to the success of evidence-based sentencing. Because risk estimation can profoundly affect defendants’ and the public’s well-being through its influence on punishment, many scholars contend that these false positives and false negatives occur too frequently for comfort (Berk et al., 2009; Berk & Bleich, 2014; Cooke & Michie, 2010; Harcourt, 2007; Hart et al., 2007).

Third, evidence-based sentencing has been accused of being destructive of individualized justice. The American criminal justice system places great value on individualized sentencing, the idea that all characteristics of each offender and offense should be considered when making sentencing decisions. As Etienne (2009) notes, this concept has been enforced by the US Supreme Court on several occasions as a means of justifying both harsher and lighter sentences (see *Williams v. New York* and *Woodson v. North Carolina*). Risk assessment instruments, however, only consider defendants unique insofar as those defendants differ on the characteristics included as factors in the risk and/or needs scores. Defendants cannot be differentiated by characteristics that are not included in risk assessment tools. Even risk assessment instruments that include a broad variety of offender characteristics cannot capture all the unique aspects of defendants and their cases that should be relevant to sentencing. This could be seen as amounting to a method of statistical profiling in which defendants are treated better or worse depending on group memberships that they often cannot change (Auerhahn, 1999; Hannah-Moffat, 2005; Starr, 2014b).

Additionally, over the course of the evidence-based sentencing process, court actors have a tendency to treat risks and needs as interchangeable concepts when the two should be considered separately (Hannah-Moffat, 1999). While risk factors signify increased likelihood of recidivism and are intended to inform risk *management* decisions, needs identify targets for treatment programs and are intended to inform risk *reduction* decisions, and ultimately, evidence-based sentencing is needed to reduce risk rather than just to manage it (Garland, 2003; Maurutto & Hannah-Moffat, 2006). Addressing defendants' criminogenic needs is thus an important goal of evidence-based sentencing. However, unsatisfied needs are often regarded by judges and other practitioners as though they are risk factors; they are used to determine how risky each defendant is but not what treatments would be useful to reduce his/her risk (Hannah-Moffat, 1999). Evidence-based sentencing thus falls short of what it has the potential to accomplish. What is needed is a broader, more flexible approach to offender evaluation that incorporates both risk identification and risk reduction (Dowden & Andrews, 2000). Such an approach would use risk and needs identification to craft a treatment plan that protects the public in the short term but also rehabilitates the offender in the longer term.

Perhaps the most strongly voiced concern with evidence-based sentencing is its potential to exacerbate extralegal disparities in sentencing outcomes (Hannah-Moffat, 2005, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b,c). Empirical evidence suggests that offender characteristics such as race, gender, and SES correlate with reoffending, but these factors are not considered legally relevant variables for sentencing (Oleson, 2011; Starr, 2014b). This leads to an interesting tension between demands for effective risk prediction and the protection of constitutional rights. On one hand, some legal scholars vehemently condemn the incorporation of these variables into risk assessment calculations; Starr (2014b) argues, for instance, that using demographic or socioeconomic variables to arrive at risk scores and sentencing decisions is a serious violation of the Equal Protection Clause, tantamount to overt discrimination. Work by Hannah-Moffat and colleagues demonstrates that risk and needs assessments are built on principles of the middle class that are both racialized and gendered (Hannah-Moffat, 2009; Hannah-Moffat & Shaw, 2001). In other words, factors that increase participants' risk scores may be "risky" characteristics for white, middle-class males but may actually not be indicative of greater recidivism or may be more commonplace among other populations. Moreover, even if demographic factors such as race, gender, and SES are not explicitly included in risk assessment instruments, they may indirectly impact risk scores and subsequent punishment through their relationship with other variables that appear in risk assessment instruments. For example, racial minority defendants have more extensive criminal histories than whites (Kutateladze et al., 2015). Because criminal histories occupy a starring role in many popular risk assessment tools, minority defendants will receive higher risk scores than white defendants because of their previous records.

On the other hand, demographic and socioeconomic characteristics of defendants, as well as variables through which those characteristics may operate indirectly, provide substantial contributions to the predictive validity of risk assessment tools. Without them, the effectiveness of the tools drops precipitously. Oleson (2011: 1396-7) ponders this issue in the context of race:

> If even something as quotidian as criminal history, a staple in traditional sentencing, can operate as a proxy for race, what variables are free from suspicion? Gender? Age? Family background? As each variable is discarded as antithetical to American legal values, the predictive value of the model dwindles until we are left with something no more robust than the best guess of a judge. (1396—7)

There's the rub. Excluding any factors that could be deemed constitutionally suspect would erode the predictive validity of risk assessment instruments, but without predictive capabilities, the instruments are of little value to the criminal justice system. If evidence-based sentencing does add to sociodemographic disparities in sentencing, adjustments will be needed to protect defendants' constitutional rights and prevent further discrimination of marginalized groups of offenders.

However, it is unclear whether or not evidence-based sentencing actually does exacerbate social inequality in sentencing (Laskorunsky, 2017). Scholars assume that risk assessment instruments would worsen disparities, but this has not been confirmed empirically (Heilbrun, 2009). The closest evidence suggesting that evidence-based sentencing may cause demographic disparities come from examinations of disparities in the predictive validity of risk assessment instruments. ProPublica, a non-profit investigative journalism organization, conducted such an evaluation on COMPAS. Using a sample of defendants from Broward County, Florida, Angwin and colleagues (2016) assessed the two-year forecasting ability of the risk assessment instrument among black and white defendants. The researchers found that though a similar number of predictive errors were made for black and white offenders, the tool was more likely to wrongly label black offenders high-risk and to wrongly label white offenders low-risk. Though the findings from this evaluation cannot speak directly to social inequality in sentencing outcomes, they do suggest that risk assessments like COMPAS may be overestimating black defendants' risk of recidivism and underestimating white defendants', errors which could generate sentencing disparities under evidence-based sentencing schemes.[1] Whiteacre (2013) similarly evaluated risk classification errors in the Level of Service Inventory-Revised (LSI-R) and found that black offenders were more often misclassified than white or Hispanic offenders, though his results were somewhat sensitive to classification cutoff decisions and the outcomes being predicted.

Due to the explicit inclusion of sociodemographic factors themselves as well as other factors that correlate with those demographic factors, the use of risk assessments at sentencing may widen punishment gaps for different race, ethnicity, gender, and SES groups. Demographic factors could directly influence risk assessment scores but could also operate indirectly through factors included in risk assessments. Conversely, risk assessments that incorporate a broad range of offender characteristics may have little to no effect on existing sentencing disparities or may even reduce them. Scholars theorize that court actors are forced to make processing and sentencing decisions under both time and resource constraints, which limits the information that can be obtained and used to inform decision-making (Albonetti, 1991; Steffensmeier et al., 1998). So it is possible

[1] It is important to acknowledge that Angwin and colleagues' (2016) findings have been heavily contested by Northpointe, the proprietary owners of the COMPAS risk assessment (Dieterich, Mendoza, & Brennan, 2016). The findings have likewise been rebutted by other researchers who argued that ProPublica's conclusions of bias are inappropriate (Flores, Bechtel, & Lowenkamp, 2016).

that risk assessments could be a mechanism through which relevant factors that would not otherwise have been considered can impact sentencing and limit the influence of race, ethnicity, gender, and SES. It is this set of possibilities that this study tests, using evidence-based sentencing based on the LSI-R.

The Level of Service Inventory-Revised

The Level of Service Inventory- Revised (LSI-R) is a popular third-generation proprietary risk assessment instrument developed by Canadian psychologists using analyses of Canadian inmates and professional judgments by Canadian probation officials (Andrews & Bonta, 1995). Originally titled the Level of Supervision Inventory (Andrews, 1982), the instrument started out as a guide for determining both appropriate levels and lengths of supervision for offenders in the correctional system. By the 1990's, the tool had been revised and renamed to reflect its new emphasis on service and treatment provisions in addition to supervision. Among states that have not adopted their own state-specific instrument, the LSI-R is the most popular risk prediction tool for use in the court system (Andrews & Bonta, 2000; Casey et al., 2014; Olver et al., 2014).

The LSI-R is based on an approach to offender assessment known as the "Risk-Needs-Responsivity" (RNR) model (Andrews et al., 1990). True to its name, the approach relies on three main principles related to the design and implementation of offender interventions. The risk principle requires that the level of service provided to each offender be matched to that offender's individual risk of reoffending. High-risk offenders should receive more service, while low-risk offenders should receive less. This principle requires a valid and reliable method of predicting offender recidivism and differentiating between high and low risk levels. The need principle specifies that offenders' criminogenic needs should also be assessed and targeted with treatment. In contrast to static risk factors, criminogenic needs are dynamic factors directly linked to criminal behavior that can be altered with intervention. With the need principle, the RNR model departs from risk management approaches by aiming to *reduce* risk rather than just to manage it (Hannah-Moffat, 2005). The third principle, responsivity, dictates that offenders should receive cognitive behavioral treatment to maximize their ability to benefit from rehabilitative interventions, and that these interventions should be individually tailored to the learning style, motivations, and strengths of offenders (Andrews et al., 1990). This principle is derived from theories of cognitive social learning, which suggest that people learn differently and will learn the most if interventions attend to their personal learning styles (Andrews & Bonta, 2006). According to the responsivity principle, treatment providers must also account for offenders' unique personal, cognitive, and social factors in order to maximize the effects of the treatment. The RNR model is highly influential and has been shown to reduce recidivism significantly when applied appropriately in rehabilitation programming (Andrews & Bonta, 2006), making it an asset to assessment instruments like the LSI-R.

The LSI-R is a 54-item assessment designed to measure offenders' propensity for future antisocial behavior. The items capture ten criminogenic domains, including criminal history, education/employment, financial characteristics, family/marital status, accommodations, leisure/recreational activity, criminal companions, alcohol/drug problems, emotional/personal attributes, and attitudes/orientations toward crime and the criminal justice system (Andrews & Bonta, 1995). Factors included in other risk assessment tools that do not appear in the LSI-R include demographic factors, current offense type, history of physical or sexual abuse, and protective factors that reduce the likelihood of future offending. Each item in the LSI-R is scored in a yes/no format, and each "yes" (meaning risk factor is present) adds one point to the risk score. Higher composite scores, which range from 0 to 54, indicate that more criminogenic risk factors are present and the offender is at a greater risk for future criminal behavior.

Many evaluations of the LSI-R have been conducted to determine its predictive validity. Generally speaking, the LSI-R performs better than chance predictions and as well or slightly better than other risk assessment instruments (Andrews & Bonta, 1995, 2006; Flores et al., 2006; Kroner & Mills, 2001). Gendreau and colleagues' (1996) meta-analysis of actuarial assessment instruments identified the LSI-R as the instrument with the highest recidivism correlation (r=.35), though only some of the comparisons achieved statistical significance. Most evaluations of the LSI-R, though, have been conducted using primarily white, male samples. Several scholars question the accuracy of the instrument for particular criminal subpopulations. A variety of studies have evaluated the predictive validity of the LSI-R for female offenders, and though female offenders sometimes receive lower composite scores than males (Coulson et al., 1996; Holsinger et al., 2006), a meta-analysis of these studies found similar effect sizes for male and female offenders, suggesting that the LSI-R is similarly predictive for males and females (Smith et al., 2009). Others have tested the efficacy of the LSI-R among racial/ethnic minority groups (Holsinger et al., 2006; Olver et al., 2014; Whiteacre, 2006), likewise finding that minority offenders earn higher scores but are subject to comparable predictive accuracy as whites.

A meta-analysis capturing the predictive accuracy of various Level of Service scales, including the LSI-R, also identified few accuracy differences across offender demographic characteristics and outcomes (Olver et al., 2014). Interestingly, the authors of this meta-analysis did note that the accuracy of the scales was largest among Canadian samples and lowest among studies conducted in the United States, a finding that likely reflects the instrument's original focus on Canadian offenders and differing crime dynamics between the U.S. and Canada.

Thus, the LSI-R can be considered an extensive, scientifically validated risk assessment tool. Substantial variation across risk assessment instruments precludes the use of any single instrument from being perfectly representative of evidence-based sentencing in the United States. Still, the LSI-R is a good instrument to use for an assessment of evidence-based sentencing because it is the most popular generic risk assessment tool and it incorporates a relatively large range of domains related to recidivism that could be sources of social inequality.

This is the context that provides fertile ground for an evaluation of disparities resulting from criminal justice policy. Evidence-based sentencing has been heralded as a potentially important mechanism for reducing incarceration rates. It involves a data-driven approach to risk evaluation, it purports not to endanger public safety while still limiting the use of prison, it turns the spotlight toward utilitarian punishment goals such as risk reduction, and it has the potential to increase objectivity and transparency in sentencing. At the same time, evidence-based sentencing is not well-understood by the public, risk evaluations are imprecise and sometimes improperly label offenders as "high risk" or "low risk", some argue that evidence-based sentencing amounts to statistical profiling, risk assessments are often used only as risk management tools when they are also intended to inform risk reduction decisions, and evidence-based sentencing has the potential to exacerbate social inequality.

The dissertation project focuses on this last concern using the LSI-R, which is a popular, well-validated risk assessment instrument that uses a wide variety of criminogenic factors to evaluate recidivism risk. The setting for the evaluation is Connecticut, a state that does not currently have an evidence-based sentencing scheme in place. The next chapter discusses Connecticut and explains why Connecticut is a suitable location for this project.

CHAPTER 4: RESEARCH CONTEXT

This study examines sentencing outcomes in a single state – Connecticut. Because it is frequently posited that individual sentencing decisions are influenced by social contexts, leading to important roles for both courtroom and community factors in sentencing disparities (Eisenstein & Jacob, 1977; Dixon, 1995; Johnson, 2005), this sentencing evaluation would be incomplete without considering the context in which the sentencing decisions are made. This chapter describes the state of Connecticut generally, provides background on the courts and the development of its sentencing system, and then discusses aspects of the penal code and other statutes that are most relevant for the present study on evidence-based sentencing and social inequality. It concludes with a commentary about the suitability of Connecticut for this evaluation of evidence-based sentencing and social inequality. To provide a time-consistent context, it describes the state of Connecticut and its criminal justice system between 2008 and 2010, rather than the present day, when possible.

Connecticut

In 2010, Connecticut had a total population of 3.57 million people. It was the third-smallest state in the country but also the fourth most densely populated (US Census Bureau, 2012). Its resident population was 77.6% white, 10.1% Black or African American, and 3.8% Asian, with the remaining percentage identifying as another race or two or more races. 13.4% also identified as Hispanic or Latino (US Census Bureau, 2010). These figures are comparable to the demographic makeup of the United States as a whole. In 2010, the US population was 72.4% white, 12.6% Black or African American, 4.8% Asian, and 16.3% Hispanic or Latino (US Census Bureau, 2010).

Connecticut is one of the wealthiest states in the US, with a high median household income and low federal poverty rate (US Census Bureau, 2017). However, it is also one of the most socioeconomically unequal. Income inequality has increased across all states since the 1970s, but the income ratio of the top 1% to the bottom 99% in Connecticut has grown substantially, swelling to 42.6 and dwarfing the total US top-to-bottom ratio of 25.3 in 2013 (Sommeiller et al., 2016). Income inequality also varies considerably across geographic regions in Connecticut; the top-to-bottom ratio ranges from 73.7 in Fairfield County, which contains four of the state's largest cities, down to Windham County, which is the least populous county in the state (Sommeiller et al., 2016).

Connecticut has long been considered a liberal state. Between 2008 and 2010, Connecticut's governor was Republican, but the state remained blue. Democrats held a firm majority in the state legislature, most of the US Senators and Representatives were Democrats, and Connecticut's Electoral College votes went to Democrat Barack Obama in the 2008 Presidential election.

Connecticut Courts and Sentencing Background

Connecticut has no county governments, so all criminal justice functions except localized police services are provided by state agencies. Superior Courts hold original jurisdiction over all criminal cases in the state. The Superior Courts are divided into 13 judicial districts ("JD's") and 20 geographical areas ("GA's"). GA courts typically handle criminal arraignments, misdemeanors, and lower-level felonies, while serious criminal matters are heard in JD courts (State of Connecticut Judicial Branch, 2017). As in other states, Connecticut has an intermediate Appellate Court and a Supreme Court that can review decisions from the lower courts.

Prior to the 1980's, Connecticut had an indeterminate sentencing system. Judges assigned a minimum and maximum term of punishment to each offender, but parole boards maintained authority over convicted offenders' served sentence lengths through discretionary release decisions. Rehabilitation was the dominant sentencing ideology. In 1981, Connecticut's General Assembly abandoned indeterminate sentencing and shifted to a determinate sentencing model in which judges assigned fixed terms of imprisonment but retained the discretion to consider a range of sentence types specified in state statutes. Along with indeterminate sentencing went discretionary parole and the emphasis on rehabilitation, replaced with a supervised home release program managed by the Department of Corrections and a new emphasis on deterrence, consistency, and retribution (Legislative Program Review & Investigations Committee, 2005). The state continued making changes throughout the 1980's and 1990's. Mandatory minimums were established for various offenses. In 1993, discretionary parole was reinstated in a limited fashion; the Connecticut Board of Parole was given jurisdiction over offenders whose sentences were at least two years (Legislative Program Review & Investigations Committee, 2005).

During this time, more offenders were sentenced to incarceration, and average prison sentences got longer. The Connecticut General Assembly implemented a truth-in-sentencing policy for violent offenders, requiring violent offenders to serve 85% of their sentences before becoming eligible for parole. Persistent offender provisions and other penalty enhancements increased the size of mandatory minimum penalties. Maximum “good time” credits were reduced by 20%, which increased time served by approximately the same percentage (Legislative Program Review & Investigations Committee, 2005). As a result, Connecticut’s prison population climbed dramatically. In 1985, only 5,422 inmates were housed in Connecticut prisons (Connecticut Sentencing Task Force, 2009). By 2010, this figure had risen to 18,416 (Connecticut Department of Correction, 2010). In conjunction with the rise in prison population, the cost of the prison system skyrocketed. Correctional expenditures hovered around $81 million in 1985 and rose to over $666 million by 2010 (Connecticut Department of Correction, 2010; Connecticut Sentencing Task Force, 2009). Adding to Connecticut’s prison woes, racial disparities in the prison population increased. By the mid-2000’s, the state was incarcerating black offenders at a rate 12 times higher than white offenders, earning it the dubious distinction of having one of the largest racial gaps in incarceration in the country (Connecticut Sentencing Task Force, 2009).

It was within this context that the state of Connecticut reevaluated the direction of its criminal justice system in the 2000’s. The Commission on Prison and Jail Overcrowding, originally created to oversee prison construction, changed course and began exploring mechanisms to reduce the prison population. In 2005, a Legislative Program Review and Investigations Committee was formed to review mandatory minimum policies, and among the Committee’s recommendations was the creation of a task force charged with assessing the state’s sentencing system and determining whether a formal Sentencing Commission was needed. This task force was created in 2006 by the Connecticut General Assembly and soon determined that Connecticut’s sentencing system was fragmented and in need of reform. It noted the overcrowding of prisons, the inefficiency and costliness of the system, and the social inequality associated with incarceration. It emphasized the need for "rational, data-driven sentencing policy" and recommended the establishment of a Connecticut Sentencing Commission that would develop "in-depth knowledge" of state sentencing practices and ultimately assist the state in "dedicating its limited resources in the most effective and efficient manner while ensuring and enhancing public safety and justice" (Connecticut Sentencing Task Force, 2009: 8). Thus, Connecticut created a permanent sentencing commission, which would become active in 2011.

In the years since the Connecticut Sentencing Commission was established, it has advanced Connecticut's sentencing process in a number of ways. One of the most recent initiatives undertaken by the Commission involves a study of evidence-based sentencing. As envisioned by the Commission, the study is intended to describe a) how closely current sentencing practices in Connecticut fit with sentences that would have been issued in an evidence-based sentencing context and b) how closely risk and needs assessment scores fit with offenders' subsequent recidivism patterns. The present study of social inequality and evidence-based sentencing is an extension of the study initially proposed by the Commission.

Connecticut currently does not use any general risk assessments to inform sentencing decisions. That said, several different risk assessment tools are used in Connecticut's adult and juvenile correctional systems. The Kingston Screening Instrument for Domestic Violence Offenders is widely used for the sentencing of domestic violence cases (Roehl & Guertin, 2000). Connecticut's Court Support Services Division (CSSD) employs the LSI-R, the Adult Substance Use Survey, and the Service

Planning Instrument for Women to evaluate risk and inform supervision and case planning decisions for Connecticut adult probationers (Millson et al., 2009). The Connecticut Board of Parole created its own Salient Factor Score, modeled off of a scale of the same name developed by the U.S. Parole Commission, to calculate inmates' likelihood of recidivating after release from prison and more recently has used the LSI-R for the same function (Ratansi & Cox, 2007). The Treatment Program Assessment Instrument is also used by the Department of Correction to assess risk of recidivism, most notably violent recidivism, among offenders sentenced to at least six months of incarceration, and the Static-99 is used to measure the risk of sexual offending. The scale captures prior correctional commitments, age, and violent behavior. For juveniles on probation, CSSD uses the Brief Risk Assessment Tool, the Juvenile Assessment Generic, and occasionally a structured interview called Assessing Individual Motivation. These instruments are used to develop probation disposition, supervision level, and planning recommendations. Since 2008, the Connecticut parole system has used the Youth Correctional Offender Management Profiling for Alternative Sanctions as its risk instrument (Kelly et al., 2013). Results from these various risk assessment instruments are not made available to sentencing judges in subsequent cases.

Risk assessments are also presently used to guide pretrial release and bail decision-making. Since the 1980's, CSSD has been tasked with making bail recommendations to minimize defendants' risk of reoffending while on bail or failing to show up for court dates. These recommendations are based on a 14-item pretrial risk assessment scale that considers factors such as marital status, living arrangements, employment, education, mental health, and past and present criminal behavior (Hedlund, 2015). Thus, Connecticut is well-versed in the incorporation of risk/needs assessments into criminal justice processing.

Connecticut Penal Code and Relevant Statutes

As in other states, criminal offenses in Connecticut are classified as either felonies or misdemeanors. However, Connecticut does not have sentencing guidelines to structure the punishments associated with these crimes. Instead, felonies, which include all offenses punishable by imprisonment for one or more years, are categorized based on offense severity as capital, class A (murder), class A, class B, class C, class D, and class E. There also a variety of unclassified (class U) felonies, which are assigned unique sentencing ranges within their statutory offense definitions. Class U felonies include arson murder, possession, sale, manufacturing, or distribution of illegal drugs, and a few firearm and weapons offenses. Misdemeanors, which include offenses punishable by imprisonment for no more than one year, are likewise divided in class A, class B, class C, class D, and class U. Each classification is associated with a set of minimum and maximum prison and probation sentences that provide some structure to sentencing decisions. These ranges are defined in Table 4.1. Compared to typical state sentencing guidelines ranges, Connecticut's penalty ranges are quite broad. To illustrate, the prison penalty range for class B felonies in Connecticut is 1-20 years, and the prison penalty range for class C felonies is 1-10 years. The widths of these ranges provide ample room for judges to use their discretion in assigning appropriate sentences.

There are a variety of additional sentencing policies, both formal and informal, that further impact judicial discretion during the sentencing process. First, Connecticut has mandatory minimum penalties in place for a variety of felonies as well as for a few misdemeanors and unclassified offenses. These penalties range from a mandatory life sentence for capital felonies (and 25 years for select Class A felonies) down to a mandatory one-year prison sentence for some Class A misdemeanors. Despite the prevalence of these penalties, a 2005 report on Connecticut mandatory minimums indicates that very few offenders are actually sentenced under mandatory minimum laws; instead, there are informal "going rates", or widely accepted penalties for specific crimes, that serve as the basis for plea negotiations and allow for the circumvention of mandatory minimum sentences (Legislative Program Review & Investigations Committee, 2005).

The effects of mandatory minimum laws are also softened by Connecticut's presumptive sentencing policy. Connecticut has two types of mandatory minimum sentences: traditional sentencing, and presumptive sentencing. In traditional mandatory minimum sentencing, judges cannot depart below the specific mandatory minimum prison term. In presumptive sentencing, the mandatory minimum penalty is the default sentence, but judges have an opportunity to impose sentences below the presumptive sentence if specific mitigating circumstances (referred to as "good cause") are present. For example, the sale of narcotics by a non-drug-dependent person carries a 5-year mandatory minimum penalty *unless* the defendant is under 18, the defendant is mentally impaired, or a judge determines that the crime was nonviolent. The majority of offenses carrying presumptive mandatory minimums are drug offenses (Legislative Program Review & Investigations Committee, 2005). A full list of both traditional mandatory minimum and presumptive sentences is available as Appendix A.

For felony offenses that do not carry a mandatory minimum sentence, judges have the discretion to suspend all or part of the incarceration sentence.[2] In effect, this enables judges to issue probation sentences for many felony offenders. Connecticut additionally has sentencing enhancements that allow judges to issue sentences above classification ranges for certain offenses. These enhancements can be tacked onto offenses for which mandatory minimums have been triggered. Offense-related enhancements include sentencing penalties for offenders who commit a crime while released on bail for a prior offense, and for offenders convicted of carjacking, terrorism, or committing a class A, B, or C felony with a firearm or assault rifle. Penalty enhancements related to offenders' criminal history are referred to as persistent offender penalties, and there are nine persistent offender categories in Connecticut's sentencing laws. These categories govern persistent dangerous felony, dangerous sexual, serious felony, serious sexual, larceny, felony, DUI felony, bigotry/bias, and assault/stalking/trespassing/threatening offenders. For a persistent offender enhancement to be added, the offender must have previously been convicted of a specific offense and incarcerated for more than a year, or the defendant's "history and character and the nature of the circumstances of the crime [must] indicate an extended period of incarceration and lifetime supervision serves the public interest" (Legislative Program Review & Investigations Committee, 2005: 14).

Connecticut's Suitability for an Assessment of Evidence-Based Sentencing

Connecticut makes for an interesting location for a study on evidence-based sentencing and social inequality. Connecticut is more liberal and correspondingly less conservative than most other states in the US. Reducing demographic disparities,

[2] This excludes Class A felony offenders, nearly all of which are subject to a mandatory minimum prison sentence.

especially racial disparities, has more consistently been a part of the liberal criminal justice reform agenda than the conservative one (Chammah, 2016; Democratic National Committee, 2017), which may mean that social disparities occupy a more prominent position in Connecticut's sentencing reform discussions than other states'. That said, despite its demographic similarities to the US as a whole, racial disparity in imprisonment and income inequality are greater in Connecticut than in most other states, which may indicate that the state's social ecology is less favorable toward marginalized groups than its political bent would suggest.

With respect to its formalized sentencing practices, Connecticut exhibits more similarities to other states in the US. The development in the 1980's of Connecticut's sentencing system into a determinate scheme with mandatory minimums, truth-in-sentencing laws, and greater sentencing structure was typical of state sentencing systems at the time; many other states made similar changes. This "tough on crime" approach led to harsh punishments in Connecticut that mirrored stiff sentences across the US. Likewise, many states have since transitioned back into primarily indeterminate sentencing schemes with active parole boards; 33 states, including Connecticut, currently run a more indeterminate sentencing system (Lawrence, 2015).

Connecticut's sentencing scheme also offers an opportunity to examine social inequality in a state without sentencing guidelines. Despite the fact that the majority of states do not have sentencing guidelines in place, most evaluations of extralegal sentencing disparities are conducted in states such as Minnesota, Ohio, and Pennsylvania that have sentencing guidelines (Griffin & Wooldredge, 2006; King & Johnson, 2016; Miethe & Moore, 1985; Steffensmeier et al., 1998). Additionally, the lack of sentencing guidelines enables this study to better capture social disparities due to individual decision-making in Connecticut's current system. If the evaluation were conducted in a state with sentencing guidelines, it would be difficult to determine whether disparities observed in the current system were attributable to individual discretion or to differences triggered by the structure of the guidelines. Comparisons between disparities in Connecticut's current sentencing system and disparities in an evidence-based system would likewise be less clear if sentencing guidelines were already in place.

Finally, Connecticut has demonstrated a strong interest in using risk and needs instruments to guide criminal justice decision-making, making it an ideal state in which to conduct the present study of evidence-based sentencing. Its familiarity with the assessment of risk and its willingness to explore the various ways in which risk assessments may be useful in criminal courts help maximize the utility of the study for policy decision-making. The study can directly inform decisions about how to increase the effectiveness and fairness of sentencing, so the findings are most immediately beneficial if the Connecticut Sentencing Commission and the Connecticut General Assembly are open to an evidence-based sentencing policy. Given that the current study was set into motion as the result of a solicitation from the Connecticut Sentencing Commission for an evaluation of the viability of evidence-based sentencing, it is clear that the present study has policy implications that are useful to the state of Connecticut.

CHAPTER 5: THEORETICAL PERSPECTIVES

Applied principles from the decision theory literature serve as a guide for understanding sentencing decisions in the courtroom. The study of decision-making is a truly interdisciplinary enterprise. People make decisions and act on them in all aspects of life, which allows decision-making to be a subject of inquiry for a variety of disciplines, including political science, psychology, economics, philosophy, sociology, and legal studies (Slovic et al., 1977). Thus, considering insights from multiple disciplines can improve our understanding of decision-making both universally and in specific contexts, such as the criminal courtroom. In this case, sociolegal theories are embedded within concepts in behavioral economics, which is itself an integration of psychological and economics traditions, to explain sentencing decisions. This chapter describes a general decision theory framework for understanding human decision-making in conditions of uncertainty, then integrates behavioral economics and sociolegal perspectives on decision theory to provide explanations for why disparities may exist in Connecticut sentencing and why the integration of actuarial risk assessments may alter them. It concludes with a statement of the research aims for this study.

Human Decision-Making and Choice Under Uncertainty: A Descriptive Framework

People make decisions every day. Some decisions are simple, with full information and only a few clear-cut alternatives, while others are more complex, with missing information and countless fuzzy options. Decision theory is concerned with how a person arrives at his or her decisions (French, 1986). It seeks to identify the attitudes, beliefs, and desires that underlie a person's choice and explain how that person used them to make it.

Central to decision theory is the consideration of choice under uncertainty. Due to its roots in economics, mathematics, and even philosophy, much of classical decision theory assumes a rational decision-maker, such that choices are made logically with full information and well-defined preferences (French, 1986). However, full rationality is not a universally reasonable assumption. In many instances, the decision-maker has some of the relevant information but is also missing some or even most of it (Simon, 1957). It is the unknown factors that are relevant to a particular decision which constitute uncertainty; they are the things a decision-maker does not know but must know in order to make a perfectly rational decision.

Uncertainty is a decidedly undesirable circumstance for decision-makers; people feel uncomfortable not knowing all information relevant to a particular decision they must make. There is some evidence that people experience greater levels of stress, both objective and subjective, when uncertainty is high (de Berker et al., 2016; Monat, Averill, & Lazarus, 1972). The idea can be demonstrated behaviorally as well; given conditions where rational subjects should be indifferent between two options in a choice set, subjects will choose the option with known risks over the option with unknown risks significantly more often than expected. This phenomenon, often referred to as ambiguity aversion, is well-documented in the empirical literature (Camerer & Weber, 1992; Ellsberg, 1961; Keren & Gerritsen, 1999; Keynes, 1921). Thus, when presented with decisions that involve unknowns, decision-makers are motivated to find ways to reduce and otherwise cope with the uncertainty.

According to Herbert Simon, in organizational contexts, some uncertainty can be eliminated through mechanisms such as divisions of labor and formal operating policies, but decision-makers will rarely be able to completely eliminate all the unknown elements relevant for a particular decision (March & Simon, 1958; Simon, 1957). When confronted with a complex decision marked by conditions of uncertainty, the best that decision-makers can achieve is a bounded rationality that takes into account incomplete information and cognitive limitations. In this state of bounded rationality, decision-making is suboptimal. Decision-makers look for “satisficing” solutions to problems, which aren’t necessarily the best, but are instead simply perceived to be good enough to achieve the intended goals.

Decision theory identifies and prescribes a variety of methods for coping with uncertainty and identifying satisficing solutions (Slovic, Fischhoff, & Lichtenstein, 1977). One concept discussed among these methods is heuristics, or rules of cognitive inference that help reduce the perceived (though not actual) complexity of a decision (Scholz, 1983). Heuristics are mental shortcuts; though they result in incomplete thinking, they also help reduce cognitive strain. Heuristics often have some level of validity, such that they can improve accuracy at least part of the time, but reliance on heuristics to make particular decisions may result in biases, or systematic errors in estimation and decision-making (Tversky & Kahneman, 1974).

As an example, one popular heuristic in prescriptive decision theory is RQP: reduce, quantify, and plug (Cohen et al., 1985; Janis & Mann, 1977). According to the general rules of RQP, when confronted with uncertainty, decision-makers should first *reduce* uncertainty by searching for as much relevant information as possible. Then they should *quantify* the magnitude of irreducible uncertainty (whatever uncertainty remains after the information search), and finally *plug* the results from the quantification into a scheme that considers uncertainty as a factor that should be used to determine the best available decision. Though this rather rigorous heuristic may be useful in some instances, its implementation is unrealistic in especially complex situations; when environmental uncertainty is very high, collecting additional information may not be feasible due to time and resource constraints, or it may reduce uncertainty by only a negligible amount. Instead of using mechanical optimization formulas to make decisions under uncertainty, Simon and other scholars suggest that decision-makers more often use simpler heuristics, ones with fewer steps that require less cognitive effort to complete (Simon, 1957; Tversky & Kahneman, 1974). These heuristics can sometimes even be triggered automatically and below the level of consciousness (Kahneman, 2011).

This is the general framework in which insights from the behavioral economics and sociolegal perspectives can be integrated to explain judicial decision-making in the absence of an evidence-based sentencing scheme. In conditions of uncertainty, decision-makers are constrained to using a bounded rationality that produces suboptimal decisions and precipitates the use of heuristics, or mental shortcuts, in order to ease the cognitive load. Because the criminal courtroom is characterized by high levels of uncertainty, it is an especially interesting setting in which to consider this perspective on decision-making. In the next section of this chapter, several specific heuristics are discussed and linked to sociolegal theorizing where appropriate to explain sentencing patterns in Connecticut.

Explaining Sentencing Disparities in Connecticut

Connecticut does not have sentencing guidelines. Instead, the hallmark of Connecticut's current sentencing structure is broad offense categories associated with wide sentencing ranges. Other legal characteristics such as criminal histories and specific circumstances of the offense undoubtedly provide context that can assist judges in determining where in the prescribed punishment range an offender should be sentenced, but the precise role that these factors play in sentencing is not formally dictated by law. Thus, judicial discretion in Connecticut's system is less constrained than judicial discretion in other states; judges have the power to assign a relatively wide variety of sentences to offenders that have similar case characteristics. This freedom, combined with the uncertainty inherent in determining appropriate sentences, creates ample room for judges to incorporate heuristics into their decision-making process.

The judgment heuristic acknowledges that people are often required to form an impression of something or someone in order to make a decision. This impression can be thought of in terms of a specific question that it answers. If, however, there is no satisfactory answer to the question that relies on little uncertainty and can be found with relative ease, decision-makers may opt instead to answer a related question that is more easily answered. The answer to the easier question then serves as a substitute for the (unknown) answer to the more complex question, saving the decision-maker some cognitive effort (Kahneman, 2011).

The concept of the judgment heuristic has straightforward applications in judicial decision-making. Sentencing judges are asked to assign appropriate sentences to offenders, and the appropriateness of each sentence is determined by a wide variety of factors that may differ from case to case. This makes the sentencing task a very complex undertaking. In order to make the best possible decision in a case, the judge must assess and then incorporate the answers to a variety of questions, including but not limited to "How much punishment does this offender deserve?" and "How likely is this offender to recidivate in the future?”. These are difficult questions to answer, as the judge has only limited information regarding the offense and the offender and cannot know exactly what will happen in the future. There is a great deal of uncertainty surrounding her decision to assign a particular sentence. To cope with this uncertainty, the judge may use the judgment heuristic to inform her impression of the appropriate sentence. Rather than answering the question “How much punishment does this offender deserve?”, for instance, the judge may instead answer the question “How much punishment does this offender *look* like he deserves?”. Rather than answering the question “Will this offender recidivate in the future?”, the judge may instead answer the question “How much does this offender *look* like a recidivist?”. These substitutions generate more readily available impressions that can then be used to inform the sentencing decision.

Two specific examples of the judgment heuristic that are particularly useful for understanding judicial discretion in a largely unregulated sentencing system are the closely related availability and representativeness heuristics. When using the availability heuristic, decision-makers evaluate the likelihood that an event has occurred or will occur by relying on the ease and speed with which examples of that event come to mind (Tversky & Kahneman, 1974). Thus, the decision-maker uses the answer to the question “Can I think of times when this event has happened before?” as a substitute for the answer to the harder question “What is the likelihood of this event happening?”. In a similar fashion, people using the representativeness heuristic evaluate the likelihood of an event being in a certain class by assessing the degree to which that event is similar to others in that class (Tversky & Kahneman, 1974). In this case, the substitute question is “Is this event similar to other events that have happened before?”.

The availability and representativeness heuristics can be readily applied in a judicial decision-making context. When sentencing judges use the judgment heuristic to avoid answering questions riddled with uncertainty, such as “Will this offender recidivate in the future?”, they may use either or both of these heuristics to subconsciously choose the questions they substitute in. Using the availability heuristic, a judge may answer whether or not the offender will recidivate by thinking of other similar offenders who have recidivated before. Similarly, a judge may use the representativeness heuristic by assessing how much the offender resembles what the judge perceives to be the typical recidivist. If the judge can easily retrieve several instances of similar offenders who recidivated, or if the offender’s characteristics align with the characteristics of what the judge would think of as a typical recidivist, the judge may believe that the offender is very likely to recidivate and issue a harsher sentence accordingly.

Predominant sociolegal perspectives on courtroom decision-making map onto this behavioral economics discussion rather seamlessly. In fact, Albonetti’s (1991) causal attribution theory begins with virtually the same premise, drawing on March and Simon’s (1958) ideas of uncertainty avoidance and bounded rationality in organizational settings (Simon, 1957). Albonetti further traces their perspective by arguing that in the courtroom, sentencing judges manage high levels of uncertainty by building ‘patterned responses’ that take into account past experiences, attitudes, and believed stereotypes (see also Clegg & Dunkerley, 1980). This stereotyping, which refers to the process of drawing inferences about individuals based on their membership in a certain group or class (Lippman, 1922), is very similar to the use of availability and representativeness heuristics, in which characteristics of a group of people or events are compared to the characteristics of a single person or event. Arguably, the stereotyping process is an example of heuristics in action, such that decision-makers who employ stereotypes are simply using cognitive rules such as "If this person has the same feature(s) as the group, I will assume this person behaves like I believe other members of the group behave."

Causal attribution theory links the idea of stereotypes with literature on causal attributions, which suggests that judgments of causality are based on personal and/or environmental factors that influence behavior (Carroll & Payne, 1976; Hawkins, 1980; Shaver, 1975). Decision-makers rely on their assessments of these factors to determine the extent to which a person's behavior is due to disposition rather than situational influences. The use of personal characteristics to infer behavioral causality can then result in decisions that consistently stereotype certain groups as being prone to a behavior. Thus, Albonetti (1991) theorizes:

> Based on the work on uncertainty avoidance and causal attribution in punishment, judges would attempt to manage uncertainty in the sentencing decision by developing "patterned responses" that are themselves the product of an attribution process influenced by causal judgments. Judges would rely on stereotypes that link race, gender, and outcomes from earlier processing stages to the likelihood of future criminal activity. (250)

This theory can explain why there may be sociodemographic disparities in Connecticut sentencing outcomes. If judges associate certain racial, ethnic, gender, and socioeconomic groups with recidivism, then those associations will be reflected in sentencing decisions that are determined in part by offenders' likelihood of recidivism. Offenders who are members of groups that judges associate with recidivism, such as racial and ethnic minorities, would be expected to receive longer sentences, on average.

Steffensmeier and colleagues' (1993; 1998) focal concerns theory expands upon Albonetti's (1991) work and further clarifies how disparities may appear in Connecticut sentencing outcomes. Rather than focusing solely on offenders' likelihood of recidivism, Steffensmeier and colleagues suggest that judges have three focal concerns when determining appropriate punishments. The first of these is blameworthiness; judges determine how deserving offenders are of punishment by evaluating both the culpability of the offender and the seriousness of the offense. The second concern is community protection. This concern encompasses offenders' likelihood of recidivism, which is the primary focus in Albonetti's (1991) causal attribution theory, but also incorporates their potential dangerousness. Finally, the third category is referred to as practical constraints. This category includes both organizational concerns, such as jail and prison capacities, and individual concerns, such as offenders' ability to serve time in prison.

Steffensmeier and colleagues (1993; 1998) argue that these three concerns guide judges' sentencing decisions, and stereotypes that relate sociodemographic factors to any of the three concerns will lead to differential punishments. In the spirit of the sociological tradition, the authors discuss the cultural context in which stereotypes linking characteristics such as race, gender, and socioeconomic status to criminality emerge. Racial minorities, they suggest, and especially young black men, are portrayed in American culture as deviant, dangerous, unstable, and drug-involved (Gibbs, 1988). Frequent depictions by media mass of young black males as violent criminals perpetuate negative stereotypes about the group in the eyes of the public (Entman, 1992; Oliver, 2003).

Similar observations can be made about stereotypes relating groups other than young black males to criminal behavior. Much like African Americans, Hispanic individuals have been associated with "innate criminality" (Holmes et al., 2008), drug use and trafficking (Curry & Corral-Comacho, 2008; Richey-Mann et al., 2006), and violence (Beckett & Sasson, 2003; Holmes et al., 2008; Welch et al., 2011). They are also often stereotyped as illegal immigrants (Curry & Corral-Comacho, 2008). Young men, due to their overrepresentation in the criminal justice system, may likewise invoke attributions of criminality and guilt (Mazzella & Feingold, 1994; Nagel & Hagan, 1983), and characteristics of low SES, such as little educational attainment and unemployment, are likewise correlated with crime and may therefore lead to attributions of criminality and criminal responsibility (Freeman, 2006; Green, 1970; Thornberry et al., 1985). These typifications make it easier for decision-makers to associate minorities, men, and low-SES individuals with factors related to sentencing such as blameworthiness, dangerousness, recidivism, and ability to serve prison time. Despite good intentions and sometimes even anti-bias training, judges may still be susceptible to allowing these stereotypes to influence their decisions.

Thus, sociolegal approaches to characterizing courtroom decision-making, such as causal attribution theory and focal concerns theory, complement behavioral economics perspectives by providing the theoretical link between sentencing disparities and decision-makers' tendency to use heuristics to make difficult choices. The combination of these approaches explains why one might observe extralegal disparities related to race/ethnicity, gender, and SES in sentencing outcomes in Connecticut.

Altering Patterns of Disparity Using Evidence-Based Sentencing

Additional theorizing is needed to explain why providing risk assessment information to judges at the time of sentencing could be expected to alter sentencing decisions and ultimately patterns of social inequality in Connecticut. First, it is important to acknowledge explicitly that judges evaluate offenders' risk of recidivism regardless of whether they are provided with actuarial risk assessments. Utilitarian perspectives on punishment emphasize the importance of linking punishment severity to offenders' likelihood of future criminal behavior (Spohn, 2009; von Hirsch, 1976). Perceptions of offender risk occupy a prominent position in sociolegal theories of punishment; causal attribution theory posits that it is attributions relating sociodemographic characteristics to risk of recidivism that lead to disadvantages at sentencing, while focal concerns theory specifies "community protection", which encompasses offenders' risk, as one of the three primary concerns judges have when issuing sentences (Albonetti, 1991; Steffensmeier et al., 1998). Interviews with judges in Pennsylvania also indicate that risk of future criminal involvement is a key consideration when punishment decisions are made (Steffensmeier et al., 1998).

Moreover, judicial perceptions of offenders' risk are likely informed in part by factors captured in actuarial risk assessments even in the absence of evidence-based sentencing. The LSI-R, for example, captures information about a variety of aspects of offender's criminal history, ranging from prior convictions and incarcerations to more unique aspects such as institutional conduct and escape history, that are available in judges' case files. Information represented in other LSI-R domains, such as employment status, residential stability, mental health status, and alcohol/drug problems, may also be known to judges at sentencing. This information may be provided in case files, but it may also be learned through questioning, or judges may be able to infer it through observation of offenders' appearance and behavior in the courtroom (Johnson & King, 2017).

The incorporation of actuarial risk assessment information into sentencing effects a change in the way that judges formulate their perceptions of offenders' risk. Even without actuarial assessments, judges conduct informal risk evaluations. Some elements of recidivism risk may be known to judges at sentencing, but there is still much uncertainty surrounding how likely each offender is to recidivate. This uncertainty invites judges to supplement their knowledge of case information with stereotypes when evaluating risk. The addition of actuarial risk information, though, should reduce judges' uncertainty about offender risk, thereby reducing judges' need to rely on stereotypes to inform their risk perceptions. In this way, introducing actuarial risk assessments may reduce the amount of sociodemographic disparities attributable to judicial decision-making; less reliance on stereotypes may translate to less of a disadvantage at sentencing for often-stereotyped groups.

At the same time, even as actuarial risk assessments decrease the use of stereotypes and reduce disparities in sentencing due to individual judicial decision-making, they may also introduce their own structural disparities into the sentencing process. Many predictors of recidivism that are captured in actuarial risk assessments also correlate with sociodemographic characteristics (Hannah-Moffat, 2005, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b), which can lead to higher risk assessment scores for certain groups. Prior research provides evidence in support of this possibility; black, Hispanic, male, and low-SES offenders receive higher average risk scores on some risk assessment instruments (Manchak et al., 2009; Monahan, Skeem, & Lowenkamp, 2017; Skeem & Lowenkamp, 2016; Skeem, Monahan, & Lowenkamp, 2016). Higher risk scores will translate into perceptions of offenders in those groups as riskier, and will lead to harsher punishment for those offenders at sentencing.

Thus, this discussion distinguishes between two potential sources of disparity in sentencing. The first source is judicial discretion; the use of stereotypes may translate into punishment disadvantages for some offenders. This source of disparity is posited to be influential particularly in Connecticut's current sentencing system, where judges are not provided with risk assessment information and uncertainty regarding offenders' risk of recidivism is high. The second potential source of disparity is structural; here disparities are not due to individual decisions made by judges, but are instead attributable to institutionalized sentencing policies that favor some offenders over others. In Connecticut, actuarial risk assessment information may be a structural source of disparity in sentencing, because certain groups of offenders will receive higher risk scores. Though the implementation of evidence-based sentencing in Connecticut may reduce disparity attributable to judicial decision-making, it may also introduce a new structural source of disparity in punishment. This dual effect makes it unclear whether net social inequality would increase or decrease under an evidence-based sentencing scheme.

Additionally, the introduction of actuarial risk assessments at sentencing has the potential to increase the impact of risk evaluation on punishment. To understand this possibility, it is useful to place risk assessment scores in the framework of what Thaler and Sunstein (2008) refer to as "choice architecture", or the design of the context in which decisions are made. As Thaler and Sunstein (2008: 3) note, "small and apparently insignificant details can have major impacts on people's behavior." Large, noticeable design elements such as the number of choices presented and small, subtle elements such as the order in which choices are presented or the decision aids that are provided are both part of a decision's choice architecture and can both influence the resultant decision (Johnson et al., 2012). Small, seemingly nonintrusive aspects of the choice architecture are referred to as "nudges" (Thaler & Sunstein). Nudges can alter people's decisions in a predictable fashion, but they do not go so far as to limit the number of available options or change decision-makers' incentives. To borrow Thaler and Sunstein's example, "Putting the fruit at eye level counts as a nudge. Banning junk food does not" (Thaler & Sunstein, 2008: 6).

Risk assessment scores can be viewed as a nudge in the choice architecture of judicial decision-making. Even if judges are not required to use the risk score information in their decision-making, the fact that they have access to a score that represents defendants' risk level makes the likelihood of recidivism more salient as a sentencing determinant. Risk scores therefore prompt judges to weight defendants' risk of recidivism more in their decision-making. When judges are presented with defendants' risk scores, they may even feel compelled to fall back on deterrence and/or incapacitation (two purposes for which risk of recidivism is informative) as the primary goal of the sentence, displacing other concerns such as retribution. The highlighting of defendants' recidivism risk, as well as the subtle emphasis on forward-looking purposes of punishment like deterrence and incapacitation, will nudge judges to alter their decision-making calculus and produce sentences that are more heavily impacted by risk evaluation.

Evidence-based sentencing has great potential to alter judicial decision-making and patterns of social inequality in Connecticut. The reduction of uncertainty in risk evaluation that risk assessments facilitate may diminish judges' reliance on stereotypes to inform sentencing while simultaneously introducing structural sources of sociodemographic disparities. Risk assessments may also increase the salience of recidivism risk and encourage judges to place more value on risk evaluation in the decision-making process. With this theorizing in mind, it is important to observe both current patterns of disparities in Connecticut and how the implementation of evidence-based sentencing would shape them.

Statement of Research Aims

The first general research aim in this study is to address the possibility that extralegal disparities exist in Connecticut's current sentencing scheme, and to assess whether or not they are related to legal case characteristics and additional factors captured in risk assessments. Taken altogether, behavioral economics principles complemented by sociolegal theorizing suggests that sociodemographic disparities in sentencing outcomes are in part the result of judges' use of heuristics and stereotyping to reduce uncertainty regarding appropriate punishments. Shackled by time constraints and incomplete information, judges make attributions about offenders based on observable characteristics such as race/ethnicity, gender, and SES and then use those attributions to inform sentencing decisions. Some groups, including racial and ethnic minorities, males, and low-SES people, are more likely to be stereotyped as possessing undesirable qualities (criminality, violence, etc.), which in turn leads to judges issuing them harsher sentences. The first research question speaks to this possibility.

RQ1: To what extent are there aggregate sociodemographic disparities in sentencing outcomes under the current sentencing scheme in Connecticut?

Aggregate, or unconditional disparities in sentencing may result from factors other than just stereotypes, however; past research tends to conclude that the bulk of the variation in sentencing outcomes is due to differences in legal factors such as offense seriousness and criminal history (Spohn, 2000; Zatz, 2000). Still, scholarship sometimes observes small but significant disparities in sentencing according to race, ethnicity, gender, and socioeconomic status even after accounting for legal and case processing factors (Baumer, 2013; Bontrager et al., 2013; D'Alessio & Stolzenberg, 1993; Spohn, 2000; Zatz, 2000). In light of these findings and the combination of behavioral economics and sociolegal theorizing, the second research question considers whether or not disparities in sentencing outcomes will still be observed after accounting for legal factors and case processing variables.

RQ2: To what extent can aggregate sociodemographic disparities be explained by legally relevant offense and case characteristics?

It is also possible that sociodemographic disparities may emerge through other indirect pathways that are not always modeled in sentencing evaluations. Scholars posit, for example, that family characteristics such as marital status influence judicial decision-making by altering perceptions of both the harshness of a prison sentence (due to separation from family) and defendants' informal social control, leading to an advantage at sentencing for married offenders as well as women (Bickle & Peterson, 1991; Daly, 1989; Freiburger, 2010). These variables are either missing or incompletely captured in much sentencing research. Additionally, the study captures several unique variables for which there has been less theorizing that may also impact sociodemographic disparities in sentencing, such as recreational activities and association with criminal companions. This study capitalizes on the wide variety of domains captured in the LSI-R risk assessment instrument to account for these factors and substantially reduce the potential for omitted variable bias that often plagues traditional sentencing research. Once the full battery of extralegal factors is considered, it is possible that sociodemographic disparities will be rendered insignificant; racial/ethnic, gender, and SES disparities may be fully explained by a combination of legal, case processing, and LSI-R risk factors. On the other hand, residual disparities may remain, as the factors included in this study cannot capture stereotyping effects that influence judges' perceptions of blameworthiness, community protection, and practical considerations.

RQ3: To what extent can observed sociodemographic disparities be explained by additional factors captured in risk assessment scores that are not often represented in sentencing evaluations?

These research questions are designed to facilitate a stepwise assessment that can first identify aggregate sociodemographic disparities before attempting to account for them using legal and extralegal factors. In this way, the first phase of the study expands on other work, such as Spohn's (2008) and Ulmer and colleagues' (2016) research, that identifies legal and extralegal mechanisms that mediate sociodemographic disproportionality in sentencing. This study advances that literature by considering a variety of defendant characteristics often identified as potentially important predictors of sentencing but rarely captured in sentencing evaluations.

The second general research aim in the study is to assess whether sentencing outcomes based on additional information from risk assessments are likely to mitigate or exacerbate sociodemographic disparities in sentencing. Scholars argue that evidence-based sentencing would generate sociodemographic (particularly racial) disparities in sentencing because the correlates of recidivism captured in risk assessment instruments such as the LSI-R are also related to defendants' demographic characteristics (Hannah-Moffat, 2005, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b). For example, even though race and ethnicity may not be an explicit consideration in the LSI-R, if black and Hispanic defendants are more likely to spend their free time engaging in unstructured activities or have primarily criminal companions, then race and ethnicity may operate indirectly through these factors, which are explicitly considered, and generate disparities in risk assessment scores. Prior research indicates that these concerns may be well-founded; a few studies observe that black, Hispanic, male, and low-SES offenders receive higher risk scores from validated risk assessment tools (Manchak et al., 2009; Monahan, Skeem, & Lowenkamp, 2017; Skeem & Lowenkamp, 2016; Skeem, Monahan, & Lowenkamp, 2016).[3] The next research question addresses this prospect.

RQ4: To what extent are there sociodemographic disparities in LSI-R composite and domain scores?

Because evidence-based sentencing entails the incorporation of risk scores into sentence decision-making, sociodemographic disparities in risk assessments like the LSI-R may translate into sociodemographic disparities in evidence-based sentencing outcomes. Other legal considerations such as offense classes, criminal records, and mandatory minimum penalties would undoubtedly continue to influence sentencing outcomes in an evidence-based system, most likely retaining their positions as the most powerful predictors, but risk scores may still shape patterns of disparity in sentencing.

[3] Results for gender differences, though, are somewhat mixed. A few studies find that women receive similar scores to men's on the LSI-R specifically (Andrews & Bonta, 2003; Lowenkamp et al., 2001).

By examining patterns of sociodemographic disparity in a simplified evidence-based sentencing scenario, this possibility can be investigated. Connecticut has not implemented an evidence-based sentencing scheme, so evidence-based sentences cannot be observed. They would represent a counterfactual to the sentences that were actually issued to offenders during the study period. However, evidence-based sentences can be simulated, which enables the researcher to offer preliminary predictions about what patterns of disparity would look like if Connecticut did incorporate the LSI-R into sentencing. The fifth research question therefore addresses whether simulated sentences based on risk assessment scores are likely to result in sociodemographic disparities.

RQ5: To what extent are there sociodemographic disparities in simulated evidence-based sentencing outcomes?

The final research question in this study addresses the oft-mentioned concern that evidence-based sentencing may actually make extralegal disparities in sentencing worse. Several scholars have argued that incorporating risk assessments into sentencing would result in harsher sentences for certain groups relative to others (Hannah-Moffat, 2005, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b). These disparities could be due to the indirect influence of membership statuses on other risk factors, such as the association between race and prior criminal history.

However, to the researcher's knowledge, the efficacy of this argument has not been assessed empirically. It is indeed possible that in Connecticut, the incorporation of risk assessments would result in wider sentencing gaps based on race, ethnicity, gender, and SES. Black, Hispanic, male, and low-SES offenders may receive even harsher sentences relative to white, female, and high-SES offenders. However, given that the empirical literature on sociodemographic disparities in non-evidence-based sentencing systems indicates that disparities exist in at least some circumstances (Baumer, 2013; Bontrager et al., 2013; Spohn, 2000), it is also possible that evidence-based sentences would result in similar or even smaller disparities. Risk assessments like the LSI-R consider such a wide range of risk factors that the impact of race, ethnicity, gender, and SES may be diluted. This study will thus inform this debate by determining whether the use of risk assessments at sentencing has the potential to exacerbate or alleviate disparities.

RQ6: To what extent do sociodemographic disparities have the potential to be larger or smaller if Connecticut transitioned to an evidence-based sentencing system?

In conclusion, this study will assess disparities in Connecticut's current sentencing system as well as consider the potential impact of a shift toward evidence-based sentencing, by evaluating disparities in a simulated evidence-based sentencing scenario. Based on theorizing and prior research, it is probable that sentencing disparities will be apparent in both Connecticut's current system and in the evidence-based scenario, but it is less clear in which system the disparities might be greater. The next chapter describes the data and methods that are used to analyze these questions.

CHAPTER 6: DATA AND METHODS

One of the several ways in which this dissertation contributes to the courts literature is by introducing a unique dataset that combines official court records with detailed information on LSI-R risk assessments administered at or near the time of sentencing. This chapter begins by describing the various types of data that are used in the project. It then discusses the dependent, independent, and control variables used in analyses before concluding with an outline of the analytic strategy that is used to evaluate the research questions laid out in the previous chapter.

Datasets

The data for this project were derived from a series of four separate datasets provided to the researcher by the Connecticut Judicial Branch Court Support Services Division (CSSD), a division of the State of Connecticut Judicial Branch that provides, among other operations, administrative services and research development.[4,5] The first dataset in this collection (hereafter referred to as "case data") constitutes a full history of court verdicts for any offender processed in a Connecticut Superior Court for a criminal offense at least once during fiscal years 2008-2010. The unit of analysis in the dataset is one case, which is defined as one verdict date for one offender.[6] Each offender

---

[4] After consulting with a representative from the Connecticut Judicial Branch, the researcher is confident that the data recorded in this study's datasets are valid and accurately reflect the characteristics of each offender, offense, sentence, and risk assessment.

[5] The researcher was provided with a total of 17 separate datasets in order to conduct the evaluation of evidence-based sentencing initially solicited by the Connecticut Sentencing Commission (see page 51). Only four of them are needed to conduct the dissertation project.

[6] The original unit of analysis in the case data was a single charge. Thus, if Offender A was given a verdict for three different charge counts on a single date, each count constituted a single observation in the dataset, and Offender A was represented in three different observations. Using the "reshape" command in STATA 14.2, the researcher reorganized the dataset to generate a new dataset in which the unit of analysis is a verdict date for a single offender. After the transformation, Offender A's three count verdicts issued on the same day are represented in a single observation. If Offender A was issued a verdict for a fourth charge on a different day, however, that fourth charge is represented in a different observation.

represented in the case data may have court cases processed prior to FY 2008 or after FY 2010 included in the dataset, but *all* offenders represented have at least one court case processed during FY 2008-FY 2010 that are included. For each observation, the case data contains information regarding charging information, offense statutes, court dispositions, and sentences issued. Court cases processed prior to FY 2008 were used to generate a count of all prior convictions in the state of Connecticut for each offender in the dataset. Cases processed after FY 2010, cases processed in juvenile court, cases that did not result in a conviction, and cases for which the most severe charge was a violation of probation, infraction, or traffic violation were also removed. This removal culled the dataset from 658,545 cases to 103,117 cases. 63,521 unique offenders are represented in the case data. The second dataset ("statute data") includes information about each criminal offense described in the Connecticut penal code. Specifically, the statute data include the name, statute number, type, class, seriousness, and general offense category for each unique offense.

The third dataset is an offender population list (hereafter referred to as "demographic data") that includes the unique identification numbers of all offenders processed through the Connecticut Judicial Branch during the study period. The demographic data also contains demographic information, specifically race/ethnicity, gender, and age, for each offender. 70,220 offenders are represented in this dataset.

The fourth dataset ("LSI-R data") is data on all LSI-R risk assessment evaluations administered to offenders by trained probation officers or other staff at the CSSD. In the LSI-R data, each observation is a single risk assessment. Offenders appear more than once in the LSI-R data if they were administered more than one risk assessment before, during, or after the study period. Risk assessments may be conducted as part of presentence investigations, or they may be conducted as intake, reassessment, or discharge evaluations during a probation sentence. The LSI-R data capture individual item scores, domain-specific subscores, and composite risk scores, as well as the date the risk assessment was administered and information about educational attainment, marital status, and employment. The dataset initially represented 207,572 assessments, but 89,198 were removed because they were administered more than one year before or after the study period, and 2,998 were removed due to missing information on one or more LSI-R item scores, leaving 115,276 eligible risk assessments in the dataset.

Data Compilation

The first time each offender is charged with a crime in a Connecticut criminal court, he/she is assigned a unique identification number catalogued in the Connecticut Case Management Information System (CMIS). This ID number follows the offender throughout subsequent Connecticut court cases, such that each offender is intended to receive only one ID number per lifetime. This ID number is included in the case data, demographic data, and LSI-R data and serves as the primary linking variable for these datasets.

Figure 6.1 shows a diagram of the data linking process. The researcher began the compilation procedure by linking case data to statute data using statute codes. In effect, this procedure added offense and severity information to the most severe conviction charge for each case in the case data.[7] In the next step, the researcher used ID numbers to link demographic data to each case represented in the case data.

The researcher then linked LSI-R data to the augmented case data. In the LSI-R data, risk assessments are associated with ID numbers but not court cases. Because a single ID number could be associated with multiple risk assessments and/or multiple court cases, the researcher used a three-step approach to match risk assessments to particular conviction dates. In the first step of this process, the researcher linked every risk assessment administered to each offender during the study period to every court case for that offender during the study period. To illustrate, if an offender received three risk assessments and was convicted on two different days during the study period, the procedure yielded six observations for that offender. In the second step of the process, assessment-conviction pairings were dropped if the assessment occurred more than one year (365 days) prior to or after the conviction. In Connecticut, risk assessments are considered valid for one year; as a result, a risk assessment administered less than one year prior to the verdict can be used as part of the pre-trial investigation and considered current (B. Sperry, personal communication, 1/28/2018). Removing assessments that occurred over one year away from their associated verdicts ensures that this study remains consistent with Connecticut's risk assessment expiration policy.[8]

---

[7] Unclassified felonies and misdemeanors were not assigned an offense type in the statutes data, so the researcher manually categorized unclassified felonies and misdemeanor offenses into the broad offense type categories used in this project. Nearly all of the unclassified offenses were drug offenses.

[8] Supplementary analysis indicates that LSI-R risk assessment scores may change over time within offenders. Using all offenders in the full LSI-R dataset who were administered more than one risk assessment (n=35,629), the researcher identified the distribution of score differences over time [difference = (score at time T2) – (score at time T1)]. Among all offenders, the mean difference is -1.00 with a standard deviation of 6.00, indicating a slight average decrease in LSI-R scores over time but substantial variation around this point estimate. Descriptive statistics partitioned by time elapsed indicate that there is slightly less variation in score differences when the assessments are less than one year apart (s.d.=5.62)

The third step in the matching process involved identifying and selecting from among instances where a single conviction date was still paired with multiple risk assessments. Some offenders received more than one risk assessment during a single sentence term, while others were administered risk assessments both before and after their conviction dates. For example, an offender could receive an intake assessment, two reassessments, and a discharge assessment for the same probation sentence. Or an offender could receive one risk assessment six months prior to their conviction date (perhaps as a discharge interview following a previous probation term) and then receive a second intake assessment for the current probation term. Because this project focuses on the influence of risk assessments administered at or near the time of sentencing, only the closest assessment linked to each conviction date was retained for analysis.

Dependent Variables

The primary outcome used in this project represents the punishment given to each defendant.[9] This variable, incarceration length, captures the number of months of incarceration ordered for each offender following a conviction.[10] For offenders who received a split sentence, this variable is measured as the length of the incarceration term, irrespective of the length of probation. For offenders who did not receive a sentence

---

than when the assessments are more than one year apart (s.d.=6.16), though variation is noteworthy in both groups.

[9] Because the case data were originally organized such that each charge was its own observation, each conviction was assigned its own sentence in the original dataset. Some offenders with multiple convictions on a single verdict date had a sentence associated with only one of those convictions (the remaining convictions had been assigned a 0 for all sentence length variables), while other offenders with multiple convictions on a single verdict date had different sentences associated with each conviction. After the case data was transformed to allow for each observation to be one offender rather than one charge, the incarceration lengths for each count on a single date were summed to produce a single incarceration sentence issued to one offender on one date. A variable identifying concurrent and consecutive sentences was used to identify and keep only the longest concurrent sentence for each verdict date.

[10] In the case data, the researcher is unable to distinguish sentences of incarceration from sentences of full time served, where the entirety of an offender's incarceration sentence was served in pretrial detention. Sentences of time served are therefore treated as incarceration sentences, and the amount of time served is recorded as the incarceration length.

involving any incarceration, this variable is measured as 0. In all analyses, incarceration encompasses both jail and prison sentences. This study's modeling strategy (discussed below) models two distinct components of this single variable: whether or not the sentence length is measured as 0, which captures the decision to incarcerate as a binary in/out outcome, and the value of the sentence length in cases where the value is greater than 0, which captures the sentence length decision.

A second dependent variable is used in one phase of the project. This variable, composite LSI-R score, is measured as an interval variable ranging from 0 to 54. It can be understood as a composite of the 10 LSI-R domain scores detailed throughout this chapter (see Appendix B). A higher score indicates a higher risk of recidivism.

Independent Variables

The primary independent variables in these analyses are several demographic characteristics often implicated for their extralegal influences on punishment outcomes. Race/ethnicity is measured as a series of dichotomous variables capturing whether the defendant is black, Hispanic, or Asian.[11] Whites are used as the reference category in all models. These racial/ethnic categories are mutually exclusive.[12] Gender is captured as a dichotomous variable in which females are measured as 0 and males are measured as 1.[13] Socioeconomic status is measured using several variables, most of which are drawn directly from the LSI-R risk assessment administered to each defendant. Though

[11] The dataset also included 66 American Indian or Alaskan native offenders, who were excluded due to their infrequent occurrence in the data.
[12] Multi-racial defendants may have unique experiences in the criminal justice system and be subject to unique treatment during sentencing, and it is possible that some of the defendants in this project's sample identify as more than one race or ethnicity. Racial/ethnic information was provided to the researcher in a mutually exclusive format, however, such that multi-racial defendants cannot be identified in these data.
[13] Race, ethnicity, and gender dummies are used as individual indicators as well as components of interaction terms capturing the effect on punishment of being a young minority male, a group which prior theorizing (see Steffensmeier et al., 1998) suggests may be particularly disadvantaged at sentencing.

socioeconomic status has been operationalized using a wide variety of factors and combinations of factors, it is generally acknowledged that socioeconomic status is determined by a combination of education, occupation, and income (Grusky, 2001). Taking advantage of the breadth of variables included in the LSI-R, this project is able to include measures that touch on all three of these socioeconomic status dimensions.

Education is captured using a set of four dichotomous variables indicating whether the defendant has completed tenth grade, has a high school degree, has completed at least some college, or has obtained some post-graduate education. Completion of less than a tenth-grade education serves as the reference category. Employment is captured using two dichotomous variables indicating whether the defendant was employed or a student full-time at the time of sentencing, and whether the defendant was employed or a student part-time at the time of sentencing. Unemployed non-students and other non-working offenders are the reference category.[14] Financial situation is captured using two variables. The first finance variable is a dichotomous measure that indicates whether the defendant was reliant upon government assistance at the time of sentencing. The second finance variable is a 4-point ordinal "problems" scale indicating whether the defendant has a satisfactory financial situation with no need for improvement (coded as 0), a relatively satisfactory financial situation with some room for improvement evident (coded as 1), a relatively unsatisfactory financial situation with a need for improvement (coded as 2), or a very unsatisfactory situation with a very clear and strong need for improvement (coded as 3).[15]

[14] The two employment measures are drawn from CSSD official records rather than from LSI-R risk assessment items.

[15] This financial satisfaction variable is more subjective than a traditional financial measure such as income and could be considered beneficial for determining each defendant's perceived placement within the social

Control Variables

Legal and Case Processing Variables. In order to provide a more detailed picture of sentencing disparities, the analyses also include several legally prescribed sentencing predictors and case processing factors that have been empirically shown to predict sentencing outcomes and may account for some portion of observed sociodemographic disparities in sentencing. Offense type is captured using five dichotomous variables indicating whether the most serious conviction offense is a violent, property, drug, sex, or weapons offense. Public order offenses are used as the reference category. Offense seriousness is measured in a series of eleven dichotomous variables that represent the offense severity classification system used by the Connecticut Judicial Branch (see p. 52, Table 4.1). Offenses may be labeled as a class A (murder), A, B, C, D, E, or U felony, or as a class A, B, C, D, or U misdemeanor.[16] Class A misdemeanors are the modal category and serve as the reference group.

In an effort to better capture the nuances of criminal histories and measure their impact on sentencing outcomes, the analyses employ two separate measures of criminal history. The first measure is a count variable indicating the number of prior adult conviction dates each offender had in Connecticut, drawn from the Connecticut Judicial Branch's administrative data. This count is right-censored at 12.[17] The second measure of criminal history is the criminal history domain score, drawn directly from each offender's

---

structure. The measure also taps into other aspects of defendants' economic well-being, including financial security, wealth, and access to resources (Sumarwan & Hira, 1993).

[16] Capital felony offenses were removed from the dataset to avoid identification, per the Connecticut Sentencing Commission's request (n=4, though only 2 had complete data and would have been included in the final dataset).

[17] Though there was no clear drop-off point in the distribution of prior convictions, censoring was needed; the maximum number of prior convictions was 161. The researcher chose 12 because this was the point at which the number of cases with each value fell below 1% of the total sample. This decision led to 4.5% of cases being censored. Sensitivity analyses indicate that the study's findings are not affected by this decision.

LSI-R risk assessment. The criminal history score ranges from 0-10 and is calculated using ten LSI-R items that capture prior and current adult convictions, juvenile arrests, incarceration history, institutional escape history, institutional misconduct, past supervision violations, and history of violence. A full list of the items used to construct the criminal history score is available in Appendix B. This score is much more comprehensive in its consideration of past behavior than the number of conviction counts and is used to complement the more traditional conviction count measure.

The number of current convictions is measured as a count variable and capped at 5.[18] To approximate the potential effects of having committed an offense while under criminal justice supervision, the analyses also incorporate a dichotomous variable indicating whether the defendant was charged with a violation of probation in addition to the primary offense.[19] Mandatory minimum application is captured using a dichotomous variable indicating whether or not the most severe offense committed is associated with a traditional or presumptive mandatory minimum. To account for potential period effects, two additional dichotomous variables indicate whether the conviction occurred in FY 2009 or FY 2010. FY 2008 serves as the reference group.

Prior literature highlights the presence of a plea discount, such that defendants who plead guilty receive lighter sentences than comparable defendants who are convicted at trial (Abrams, 2011b; King et al., 2005). In this study, the researcher is unable to directly capture mode of conviction. However, these analyses do use three measures to

---

[18] Without censoring, the number of current convictions ranged from 1 to 31. As was the case for prior convictions, the cap for current convictions was set at 5 because this was the point at which the number of cases with each value fell below 1% of the total sample. This decision led to 1.8% of the sample being censored. Sensitivity analyses indicate that the study's findings are not affected by this decision.

[19] Probation violations are included in the analysis only if they constitute new offenses; violations of probation conditions are excluded. The researcher is unable to capture parole violations in the available datasets.

approximate plea bargaining. Count bargaining is measured as a dichotomous variable indicating whether defendants received fewer conviction counts than charge counts. If any counts were dropped between charging and conviction, this variable is labeled 1; otherwise, the variable is counted as 0. Charge bargaining is measured using two variables. First, charge reductions are measured as a dichotomous variable indicating whether defendants' highest conviction offense falls into a lower offense severity category than their highest charged offense. If the highest conviction offense severity is lower than the highest charge offense severity based on Connecticut's offense severity classification, the variable is labeled 1; otherwise, the variable receives a value of 0. Second, to capture plea deals in which mandatory minimums are avoided, a dichotomous variable will indicate whether or not defendants' filing charges involved a mandatory minimum but their conviction charges did not. For offenders who were charged with a mandatory minimum offense but were not convicted of one, this variable is labeled 1. For all other offenders, this variable is measured as 0.[20]

Risk Factors. In addition to the legal and case processing variables, a wide range of other defendant risk factors are included as sentencing predictors. Age at sentencing is represented using a set of categorical variables indicating whether the offender is age 18-24, 25-34, 35-44, or 45+. 18-24 year-old offenders will be the reference group. This ordinal measure of age is used because it allows for nonlinear relationships to characterize age disparities (Steffensmeier et al., 1995). The inclusion of several domain

[20] It is possible that charges may be reduced or counts/mandatory minimum may be dropped due to circumstances other than plea deals; for instance, prosecutors may drop charges of their own accord or judges may throw out charges due to insufficient evidence. There are also other forms of plea bargaining, such as fact bargaining and sentence bargaining, that are not captured in these measures, and defendants may plead guilty without obtaining any concessions in return. Still, the prominence of count and especially charge bargaining are well-acknowledged in the empirical literature (Ball, 2006; Piehl & Bushway, 2007, Shermer & Johnson, 2010), and it is expected that these measures will together create the best approximation of plea deals possible in this study.

scores drawn from the LSI-R risk assessment makes the evaluation of Connecticut's current sentencing scheme particularly unique, because these scores introduce defendant characteristics that are very rarely available in sentencing evaluations but may play a pivotal role in explaining sociodemographic disparities. Domain scores are constructed using two or more LSI-R items. Each item contributes either 0 points or 1 point to the domain score; thus, the minimum domain score is 0 and the maximum domain score is equal to the number of items in that domain. A higher domain score indicates a less satisfactory or less desirable situation.

A full list of the items used in the calculation of each domain in the LSI-R is available as Appendix B.[21] The family and marital domain is comprised of four items that cover dissatisfaction with marital situation, non-rewarding parental and other relative situations, and criminal family members. Because the family and marital domain of the LSI-R does not include an explicit indicator of marital status (but marital status is recorded at the time of the assessment), a supplementary dichotomous variable indicates whether or not the offender is married.

The accommodation domain is comprised of three items that cover living arrangement satisfaction, residential stability, and whether defendants live in a high crime neighborhood. The leisure/recreation domain is comprised of two items that cover participation in organized activities and whether or not defendants could make better use of their time. The companions domain is comprised of five items that capture social isolation as well as how many criminal and anti-criminal friends and acquaintances the

[21] Two of the ten risk domains are discussed above; the criminal history domain score is grouped with other legal factors, and the two questions that comprise the financial domain are both used as indicators of SES. Additionally, because education and employment are both used as indicators of SES, the Education/Employment risk domain is also excluded from analysis.

defendant has. The alcohol and drug domain is comprised of nine items that cover past and present alcohol/drug problems, family alcohol/drug problems, law violations and interference with school/work, and medical treatment for drug/alcohol problems. The emotional/personal domain is comprised of five items that cover mental health interference, past and present treatments, and psychological assessments. The attitudes/orientations domain is comprised of four items that cover defendants’ attitudes toward crime, toward convention, toward their own sentences, and toward criminal justice supervision.

Finally, many offenders were convicted in a Connecticut superior court more than once during the study period, which means they appear more than once in the dataset. To capture the effect of repeat appearances in the dataset (which could be construed as a post-hoc indicator of risk), a dichotomous variable indicates whether the offender has a previous case included in the dataset. This variable is recorded as 0 if the case is the offender’s first appearance during the study period, and 1 otherwise.

Analytical Strategy

Phase 1. The analyses for this project are divided into three phases. In Phase 1, three-part stepwise regression analyses are used to evaluate sociodemographic disparities in Connecticut’s current sentencing scheme. For all analyses, the in/out decision is modeled using the logistic regression component of a negative binomial hurdle model, and incarceration length is modeled using the count component (a zero-truncated negative binomial regression) of the same hurdle model.

Criminologists have grappled with the issue of how to properly model incarceration length for decades. The assumptions required for an ordinary least squares

(OLS) regression are not typically met in incarceration sentence data, which are often characterized by an extreme positive skew, heteroscedasticity in which the variance increases with the mean, and a bound at 0. The application of OLS to sentencing data can therefore lead to biased estimates. Log-transforming sentence lengths in order to account for the distribution's skew and heteroscedasticity before running OLS is a common practice (e.g. Fischman & Schanzenbach, 2012; Spohn & DeLone, 2000; Steffensmeier et al., 1998), but there is some evidence that this strategy still results in biased estimates in the presence of severe heteroscedasticity (Hilbe, 2014; Santos Silva & Tenreyo, 2006).

It is also important to consider the potential for selection bias into incarceration in sentence length evaluations (Berk, 1983), and many scholars have used two-stage models with a Heckman correction to address this possibility (Steffensmeier & Demuth, 2001; Ulmer & Johnson, 2004). However, execution errors such as the frequent failure to incorporate an exclusion restriction, which specifies at least one variable that affects selection into incarceration but does not affect incarceration length, are highly problematic and may outweigh any benefits obtained from the model (Bushway, Johnson, & Slocum, 2007). The Tobit regression model, which treats 0's as evidence of censored observations, is another model that has been used to address this selection concern (Helms & Jacobs, 2002; Kurlychek & Johnson, 2004). However, the Tobit model assumes that censored observations have "true" but unobservable values at or below the censoring point. In the context of sentencing, this assumption would suggest that many cases receiving 0 days of incarceration have true values below 0, which is conceptually problematic. Moreover, the Tobit model assumes that the same underlying causal processes determine both the censoring process (i.e. the incarceration length decision) and sentence length decisions, which may not be the case (Steffensmeier & Demuth, 2001).

Hester and Hartman (2017) conceptualize sentence lengths as event counts, capturing the number of months an offender is sentenced to incarceration. Just like event counts, sentence lengths are positive values, bounded at 0, with a variance that increases with the mean. Because incarceration sentences can be considered a count variable with a large number of 0's, Hester and Hartman argue that mixture models such as the hurdle and zero-inflated models can be useful for modeling them. Both hurdle and zero-inflated models account for excess zeros by linking a binary model with a count model. The zero-inflated negative binomial (ZINB) model assumes there are two distinct processes that can lead to defendants receiving an outcome value (in this case, an incarceration sentence value) of 0. First, some observations may be ineligible for a positive outcome value. These observations can be thought of as "certain zeros"- they would not under any discretionary conditions be assigned a positive value. In the second process, observations are eligible for a positive value, but they may be given a 0 anyway. The ZINB accounts for these two disparate processes by first modeling the likelihood that each offender is a "certain zero" with a binary model (the "inflation" logit or probit model), then separately predicting the sentence length for offenders who are not certain zeros with a negative binomial model (the "count" model).

On the other hand, the hurdle model assumes there is a single threshold or "hurdle" that must be overcome in order for an observation to obtain a positive count. In the case of sentencing data, this translates to a threshold that must be met in order for a defendant to be sentenced to incarceration and assigned a sentence length. A binary model (a logit or probit regression) predicts observations that have zero counts (i.e. are not sentenced to incarceration), and then a zero-truncated count model predicts the outcome (sentence length) for observations that have met the threshold. The two models are connected using a log-likelihood that is the sum of the logs of the probabilities that the outcome is 0, 1, and a positive count (Hilbe, 2014). As Hester & Hartman note, this modeling strategy is consistent with prior theorizing on judicial decision-making that suggests sentencing is a two-step process: first judges determine whether or not to incarcerate each offender, then they determine the sentence length for offenders whom they feel need to be incarcerated (e.g. Spohn & Cederblom, 1991). As Hester and Hartman (2017) also mention, the estimates derived from hurdle and zero-inflated models are often very similar, but the assumptions underlying each model are quite different and must be considered carefully before one is selected for analysis.

For this study's analyses, the researcher argues that the single threshold assumption required to model an outcome using a negative binomial hurdle model is reasonable. All cases in the sample are legally eligible for an incarceration sentence; even offenses in the lowest offense severity class, Class D misdemeanors, can be penalized with one month of incarceration without any required justification from the sentencing judge. Therefore, in all cases where an alternative sanction is issued, the sentence can be interpreted as the result of judicial discretion. The judge chose to grant leniency and not issue an incarceration sentence. There is plenty of evidence for this process; within every individual offense class or prior record group, some offenders were incarcerated and some were not. Because there is a single decision-making processes through which an offender can be assigned a non-incarceration sentence, then, the negative binomial hurdle model is appropriate for this study's analyses.[22]

In this study's analysis, the jointly captured in/out and incarceration length decisions are regressed first on the primary independent variables, which capture defendants' race, ethnicity, gender, and socioeconomic status. This first model does not include any control variables. Because the data account for cases processed in courtrooms all over the state of Connecticut, and cases that go through one courtroom may be more similar to each other than they are to cases in other courtrooms, the analysis also adjusts for correlated error by clustering cases by the courtroom in which they were processed. The coefficients for each of the key predictors in this first model represent naïve aggregate sociodemographic disparities in sentencing outcomes (RQ1).

In the next step of the stepwise regression analyses, the researcher builds upon the disparity model by adding legal variables and case processing factors as predictors.[23] These additions determine how much sociodemographic disparity in the in/out and sentence length decisions can be accounted for by these characteristics (RQ2). In the final step of the stepwise regressions, the risk domain scores drawn from the LSI-R risk assessment are also added to the model. These risk domain scores, which represent a range of unique offender characteristics, have the potential to greatly improve the explanatory power of the sentencing model and to address how much of observed sociodemographic disparities are due to differences in seldom-captured risk factors such

[22] Supplemental analyses [results not shown] indicate that using a zero-inflated negative binomial model rather than a hurdle model does not substantially alter conclusions in this study.

[23] Because all Class A felony murders are subject to a traditional mandatory minimum incarceration sentence, the binary indicator of felony murder perfectly predicts the decision to incarcerate in the logistic regression component of the hurdle model. This indicator and the eight felony murder cases are therefore dropped from the logistic regression but are included in the zero-truncated negative binomial portion of the hurdle model.

as residential stability, leisure activities, criminal companions, alcohol/drug problems, and attitudes toward the criminal justice system (RQ3).

Phase 2. The second phase of the project evaluates disparities in LSI-R scores and then constructs and evaluates disparities in evidence-based sentences. First, descriptive statistics demonstrate whether or not there are unconditional differences in LSI-R composite and domain scores for each racial/ethnic, gender, education, employment, and financial group. Regression analyses, specifically ordinary least squares regression and ordered logistic regression, then assess how much of these disparities remain once other sociodemographic characteristics have been accounted for (RQ4).[24] This process shows the magnitude of the disparities in LSI-R composite scores and identifies the domains that contribute most to those overall disparities.

In order to provide some indication of the extent to which disparities in LSI-R scores may translate into disparities in evidence-based punishment, the rest of Phase 2 simulates sentences in a simple evidence-based scenario and evaluates disparities in them. Each offender in the study sample is assigned a simulated evidence-based sentence based solely on legal characteristics, case processing factors, and LSI-R composite scores. To understand how these simulated sentences are determined, it is useful to first consider two assumptions on which the construction procedure used in this study relies.

The first assumption is that each offender processed in a Connecticut superior court has a single latent risk of recidivism that is approximated by judges at the time of

[24] The distribution of LSI-R composite risk scores approximates a normal distribution, with a mean of 26.2, a slight left skew (skewness=-.25), and a slightly flat curvature (kurtosis=2.75). The researcher therefore finds ordinary least squares regression to be appropriate for this outcome. On the other hand, though individual domains may have as many as 10 unique values within them, none of the domain distributions achieve normality, so ordinary least squares is inappropriate and the researcher uses ordered logistic regressions to model these outcomes.

sentencing and by scores on the LSI-R, albeit with different levels of precision. Though the LSI-R may incorporate more risk factors into its scoring than judges incorporate into their risk perceptions, both evaluations tap into the same underlying risk characteristic. Under this assumption, LSI-R scores can be used as a proxy for judges' risk evaluation in the absence of evidence-based sentencing. In Connecticut's current system, where offenders are administered LSI-R risk assessments but judges are not provided with LSI-R information at the time of sentencing, the relationship between LSI-R scores and punishment approximates the relationship between offender risk and punishment.

The second assumption is that the integration of actuarial risk assessments into Connecticut sentencing would not alter the role that a) legal/case processing characteristics and b) risk evaluation play in the formation of punishment decisions. In other words, the introduction of actuarial risk assessments into sentencing would affect punishment only by changing the precision of judges' risk perceptions. It would not change how much weight is given to legal and case processing characteristics or to risk of recidivism in the judicial decision-making process. With this assumption in place, estimates of the relationship between legal/case processing factors and punishment, and the relationship between offender risk and punishment, in Connecticut's current sentencing system can be extrapolated to relationships among those same variables in the simulated evidence-based sentencing scenario.

These two assumptions may be tenuous. First, it is possible that judges conceptualize offender risk differently than the LSI-R captures it, which would mean that the two appraisals are not necessarily tapping into the same latent characteristic. Furthermore, as discussed in Chapter 5, the introduction of actuarial risk assessments at sentencing is expected to strengthen the role that risk plays in determining appropriate sentences. Assessment scores may serve as "nudges" that increase the salience of risk evaluation and encourage judges to weight offender risk more heavily and other factors less heavily in their punishment decisions.

Therefore, it is important to consider how simulated patterns of disparity should be different if these assumptions are not met. If the first assumption does not hold, and judges' risk evaluations do not estimate the same latent risk characteristic that the LSI-R does, the observed relationship between LSI-R scores and punishment in Connecticut's current system is likely to underestimate the strength of the relationship between recidivism risk and punishment, biasing it toward 0. Elements of perceived risk that judges consider but that the LSI-R does not will not be captured in the risk-punishment relationship, and the observed relationship will not be as strong as it should be. The relationship between risk and punishment in Connecticut's current system will be estimated conservatively.

If the second assumption does not hold, which is possible because evidence-based sentencing is expected to increase the impact of risk on punishment decisions, then using the relationship between risk and punishment in Connecticut's current sentencing scheme to approximate the same relationship in a simulated evidence-based scenario will result in an underestimation of the risk effect in simulated evidence-based sentences. The introduction of risk assessments may increase the salience and importance of risk at sentencing, but this study's modeling strategy is unable to account for that possibility. The relationship between risk and punishment will be weaker in the simulated evidence-based scenario than it would be in a true evidence-based sentencing scheme. Moreover, disparities in the simulated sentences are dependent upon the risk-punishment relationship, which means that if the risk-punishment relationship in evidence-based sentencing is estimated conservatively, disparities in the simulated evidence-based scenario are likely to be conservative as well.

With these assumptions in mind, the first step in creating an evidence-based sentence scenario is to estimate the effects of legal/case processing factors and risk on punishment in Connecticut. Incarceration lengths in Connecticut's current system, for which all non-incarceration sentences are assigned a value of 0, are first regressed on all legal and case processing variables as well as composite LSI-R risk assessment scores using a negative binomial hurdle model. This model yields estimates of the associations between legal/case processing factors and punishment and LSI-R scores and punishment in Connecticut's current system. As described above, the relationship between LSI-R scores and punishment can be treated as a proxy for the relationship between risk and punishment even when judges are not provided with LSI-R scores at sentencing.

Because the study also assumes that the strength of these associations will not change if risk assessments are incorporated into sentencing, they are used to generate an evidence-based sentencing scenario in Connecticut. Using post-estimation commands, the researcher derives predicted incarceration probabilities from the model that can be used to assign each case a probability of incarceration. These probabilities represent the likelihood that each offender would be assigned an incarceration sentence if the only determinants of punishment were legal/ case processing factors and risk of recidivism, as captured by the LSI-R. This set of conditions results in simulated sentences that do not incorporate judicial discretion. If judges were to weight risk scores more heavily in their decision-making for certain types of offenders (e.g. for minority offenders), this would not be reflected in the simulated evidence-based sentencing scenario.

Using the predicted probabilities, each offender is hard-classified as receiving either an incarceration or non-incarceration sentence in the evidence-based scenario. Hard classification employs a threshold rule; all offenders who are assigned a probability of incarceration that is greater than .5 are assigned incarceration, while all offenders who receive a probability of incarceration that is less than or equal to .5 are assigned probation. This ensures that offenders who are statistically more likely to be incarcerated than to be sentenced to an alternative sanction are assigned evidence-based incarceration, and vice versa.[25] Thresholds of .4 and .6 are used as alternative specifications in the analyses, to evaluate the effects of the threshold decision on the study's findings.

After every offender has been assigned an incarceration or non-incarceration sentence in the simulated evidence-based scenario, the researcher uses a second post-estimation procedure to generate a predicted incarceration length for each case. Offenders who were assigned incarceration are then ascribed their predicted sentence length. Predicted sentence lengths are recorded in a new variable, evidence-based incarceration length. Offenders who were assigned an alternative sanction are ascribed a 0 as their evidence-based incarceration length.

---

[25] Because non-incarceration sentences are substantially more common than incarceration sentences in this sample, it is expected that most groups will be assigned a <.5 probability. This decision rule will likely result in a greater proportion of offenders being assigned alternative sanction sentences than are observed in the actual case data. However, the researcher argues that this consequence may serve as an advantage and increase the utility of the project. Evidence-based sentencing is touted as a method for reducing the use of incarceration by diverting offenders away from incarceration, and evidence indicates that states that have already implemented evidence-based sentencing do use it to reduce incarceration (Marcus, 2009; Warren, 2007). By assigning more alternative sanction sentences in the evidence-based scenario, the project models evidence-based sentencing in a way that moves toward achieving this goal.

Though mandatory minimum penalties may not be used frequently, they may have meaningful impacts on sentencing outcomes for offenders who are subject to them. To account for mandatory minimum sentencing in the evidence-based scenario, all offenders whose highest conviction charge is subject to a mandatory minimum penalty have their evidence-based incarceration length raised to meet the mandatory minimum. For example, if an offender received an initial evidence-based sentence of probation but was subject to a mandatory minimum penalty of one year of incarceration, the offender's predicted evidence-based sentence is bumped up to one year of incarceration for the analysis. Offenders subject to a mandatory minimum whose evidence-based incarceration lengths are already greater than the mandatory minimum penalty do not have their evidence-based sentences altered.

This procedure yields a set of simulated evidence-based sentences based on legal/case processing factors and LSI-R risk scores. Again, it is important to note that this simulation procedure entirely ignores the role of judicial discretion in determining punishment; in the evidence-based scenario, only the prescribed legal, case processing, and risk factors affect sentencing outcomes. Because discretion is such an integral part of decision-making throughout the criminal justice system, however, the evidence-based scenario depicted in this study is not a realistic one. It is extremely unlikely that judges would cede all individual discretion in response to the introduction of actuarial risk assessments at sentencing. Therefore, the evidence-based scenario in the study cannot be treated as an accurate counterfactual to the observed sentences issued to offenders in the study's sample. If Connecticut were to implement an evidence-based sentencing scheme, judges would exercise their discretion in individual cases, and the resultant evidence-based sentences would fail to map perfectly onto the sentences assigned in this study's evidence-based sentencing scenario. In short, there is error in the construction of the evidence-based scenario.

Rather than representing actual evidence-based sentences, the evidence-based scenario can instead be used to demonstrate how LSI-R scores can be expected to impact punishment outcomes independent of judicial discretion. This information is useful for understanding how differences in risk scores may translate into differences in punishment and punishment disparities. With this idea in mind, the outcomes generated in the evidence-based sentencing scenario in this study can be assessed for sociodemographic disparities. To establish the magnitude of disparities in evidence-based sentences, the evidence-based in/out decision and evidence-based incarceration length are regressed on race, ethnicity, gender, and SES variables, again using a negative binomial hurdle model. The coefficients in this model represent aggregate sociodemographic disparities in simulated evidence-based scenario (RQ5).

Phase 3. In the final phase of the project, the aggregate sociodemographic disparities in actual sentences identified in Phase 1 are compared to the aggregate sociodemographic disparities in the evidence-based scenario simulated in Phase 2. Because the two outcomes, actual sentences and evidence-based sentences, are recorded for the same sample of offenders, it is likely that errors in the two models will be correlated. The researcher therefore uses seemingly unrelated estimation to combine the estimation results from the two hurdle models, allowing for unbiased comparisons between models.[26] Chi-square tests are used to test for the equality of coefficients across

[26] Seemingly unrelated estimation ("suest" in Stata) combines the estimation results from two separate equations into a single vector of parameter estimates and a variance/covariance matrix. Because the

models for each independent variable in the logit and count components of the hurdle models. To illustrate, the effect of being black on the probability of receiving an actual incarceration sentence is tested against the effect of being black on the probability of receiving a simulated evidence-based incarceration sentence. Likewise, the effect of being black on actual incarceration length is tested against the effect of being black on simulated evidence-based sentence length. These tests address whether sociodemographic disparities in the simplified evidence-based scenario are larger, smaller, or approximately equal to disparities in Connecticut's current sentencing system.

(co)variance matrix is estimated using both models simultaneously, it accounts for correlated error terms between the two models and allows the user to conduct unbiased tests for cross-model hypotheses. Note that because the equations were not estimated simultaneously in a single system, this procedure does not increase efficiency. Its advantage in this analysis is its allowance for unbiased tests of equality across coefficients from different models.

CHAPTER 7: DATA COMPILATION & DESCRIPTIVES

This chapter first details match rates resulting from the linking procedures used to connect case data with offense statute information, offender demographics, and risk assessment scores. It then describes the sample using descriptive statistics related to independent, dependent, and control variables, and considers the extent of sample selection resulting from the linking procedures.

Dataset Compilation Results

Dataset match rates are listed in Table 7.1. Each row represents one step in the matching process. The first step in the data compilation process links case data to statute data. This yields a 100% match rate for the case data; every conviction charge included in the case data was assigned offense and severity information.

Demographic data is then added to the case data using ID numbers. Again, 100% of the cases represented in the case data are matched to offender demographic information in the demographic data. However, only 90% of the offenders represented in the demographic data are matched to any case data. Of the 6,699 offenders represented in the demographic data who are not matched to any case data, 2,384 (36%) have cases in the study period that were removed because they were processed in juvenile court, did not result in a conviction, or had a most severe charge of a capital felony, violation of probation, infraction, or traffic violation. An additional 2,239 (33%) only had cases outside the study period, and the remaining 2,076 (31%) cannot be linked to any conviction information inside or outside the study period. Offenders in the latter two groups are likely unmatched because of CSSD's data extraction procedure; the case data received by the researcher was created by extracting and pairing information from two separate databases at the CSSD, and a failure to pair results in case data that is missing all offender, charge, and verdict information (B. Sperry, personal communication, 1/25/2018). The case data received by the researcher contains 7,023 such charges. It is unclear whether the failure to pair systematically affected any particular types of charges or cases.

The third data merge links LSI-R risk assessment information to the case data. As explained on pages 74-75, this linking procedure has three steps: offender ID numbers are used to link every risk assessment administered to one offender to every court case that offender had during the study period, then case-risk assessment pairings with more than a one-year difference between the assessment and conviction dates are removed, then for conviction dates still associated with multiple risk assessments, only the case-risk assessment pairing with the shortest time between the conviction and assessment dates is retained. This procedure yields a full dataset in which each case is represented once, but one risk assessment can be paired with multiple cases.

Unlike the prior linking procedures, connecting LSI-R data to the augmented case data does not yield a 100% match rate for offenders in the case data; 35,591 cases (35%) representing 19,912 offenders are not ultimately matched to a risk assessment and are therefore excluded from all analysis.[27] Given the procedure used to link risk assessments to particular cases, failures in this match are due to two reasons. First, 22,124 cases representing 15,330 offenders are not paired with any risk assessments from the LSI-R

[27] Attempts to use other risk assessments conducted by the Connecticut DOC in place of LSI-R scores in cases where offenders are not matched to a completed LSI-R risk assessment yield only 2,400 additional matched cases. The researcher opts not to use this strategy to increase the sample size and reduce any selection effect because a) DOC risk scores are captured on a 0-5 scale and do not allow the same level of precision in risk calculation that the LSI-R does, and b) DOC risk assessments do not provide information about any dimensions of SES, which means that all cases with a DOC risk assessment but no LSI-R assessment are be missing information on all SES indicators.

data, meaning that 15,330 offenders do not have complete information on any risk assessment conducted during or within one year of the study period. Second, 13,467 cases representing 4,582 offenders are removed from the dataset because the offenders in those cases were only administered a risk assessment more than one year away from the conviction date.

This 65% match rate is not entirely surprising; CSSD began introducing key performance indicators, including risk assessment timeliness measures, in 2010 and found in that year that initial compliance with the risk assessment policy was in the "mid-70% range" (B. Sperry, personal communication, 1/11/2018). Moreover, CSSD administered the LSI-R to offenders as part of either pre-sentence investigations or intake/reassessment/discharge interviews for probation sentences, not as part of intake/reassessment interviews for incarceration sentences. Offenders assigned lengthy prison sentences, then, should be less likely to have received an LSI-R risk assessment within one year of their conviction date.

Following the data linking process, there are 67,526 cases in the case data with complete risk assessment information. However, data on an independent or control variable in the case data is missing for 474 cases, and there are only 66 cases with Native American/Alaskan native offenders in the dataset. Additionally, as described in Footnote 16, there are 2 capital felony cases in the dataset with complete LSI-R data. These cases are all excluded from the dataset, resulting in a final sample size of 66,984 cases representing 43,165 offenders. This sample is used for all analyses.

Descriptive Statistics

Descriptive statistics for the final sample are detailed in Table 7.2. One-third of the cases in the final sample resulted in an incarceration sentence, and the average incarceration sentence length is just under 20 months. Drug and public order offenses are the most frequent offense types, with 30% each. Property offenses comprise 22% of the sample, violent offenses account for 15%, and sex and weapons offenses together comprise only 3%. Most of the offenses are misdemeanors (62% vs 38%), and the largest single offense severity class is Class A misdemeanors, which account for 34% of the sample by themselves. Nearly all of the other misdemeanors are Class B and C misdemeanors, while Class D and unclassified misdemeanors are quite rare. Among the felony classes, Class D and unclassified felonies occur most frequently, at 14-15% each of the total sample. Class B & C felonies account for 3% and 5% each, while Class A and Class E felonies, and Class A murders, are also very rare.

The average number of prior adult convictions in the sample is 3.28, and the average criminal history domain score for offenders in the sample is 4.73 out of 10. The average case has between 1 and 2 conviction counts. 15% of the cases in the sample involved a violation of probation charge alongside the primary offense. Consistent with the Connecticut Legislative Program Review & Investigations Committee's (2005) report, which suggested that mandatory minimum sentences are applied very rarely, only 1% of primary conviction charges received a mandatory minimum penalty. The study's three measures of plea bargaining occur at very different rates; count reductions between charging and conviction occur in 57% of cases, offense severity reductions occur half as often, in 28% of cases, and only 2% of cases involved a mandatory minimum filing charge but no mandatory minimum conviction. The distribution of cases is fairly equal across fiscal years, though FY 2010 is slightly underrepresented.

Moving to offender demographics and risk factors, half of the cases in the final sample involve a white offender. Another 30% are black, while 20% are Hispanic and less than 1% are Asian. 17% of the final sample offenders are female. Most of the sample has completed 10th grade or has a high school degree (68% combined), while another 16% each completed either less than 10th grade or some college. Offenders with some post-graduate education comprise only 1% of the sample. 62% of the sample cases involve an unemployed offender, about one half involve an offender who receives financial assistance, and the average score on the financial problems scale is approximately a 2, which corresponds with "a relatively unsatisfactory financial situation with a need for improvement".

It does not appear that multicollinearity among the SES components used in this study, education, employment, and financial situation, is likely to be a substantial problem in the analyses. Table 7.3 shows Goodman and Kruskal's gamma coefficients for pairs of SES components. The Goodman and Kruskal's gamma coefficient measures the strength of association between two ordinal variables. Values range from -1, a perfect negative association, to 1, a perfect positive association (Goodman & Kruskal, 1954). As Table 7.3 shows, the strongest associations are among employment and the financial variables, with gammas in the vicinity of .50 that indicate only moderate associations. Education appears to have weaker associations with all other components.

The directions of all associations suggest that indicators of high SES tend to accompany other indicators of high SES, and indicators of low SES tend to accompany other indicators of low SES. Higher education is associated with more employment, a lower likelihood of receiving financial assistance, and fewer financial problems. More employment is associated with a lower likelihood of receiving financial assistance and fewer financial problems. And finally, receiving financial assistance is associated with more financial problems.

Turning back to Table 7.2, the age distribution slants toward younger offenders; about one-third are under 25 years old, and each older age group represents progressively less of the sample. 10% of the sample is married. The average LSI-R domain scores vary substantially across domains. For the family/marital domain, the average score is a 2.04 out of 4, which translates to 51% of the total points possible. The alcohol/drug and companions domains also average near the center of their possible ranges, at 4.26 out of 9 (47%) and 2.74 out of 5 (55%) respectively. Offenders tend to score at the lower ends of the accommodations domain, with an average of .99 out of 3 (33%), the emotional/personal domain, with an average 1.70 out of 5 (34%), and the attitudes/orientations domain, with an average of 1.38 out of 4 (35%). Finally, offenders tend toward the upper end of the leisure/recreation domain, averaging a 1.60 out of 2 (80%).

To gauge the impact of sample selection due to match failures between the case data and LSI-R data (as well as missing data), Table 7.4 compares demographic and case information for the final sample and for the 36,133 cases that were either unmatched to an LSI-R assessment or missing data on one or more variables. This comparison yields several noticeable differences. As expected, cases with longer incarceration sentences are more likely to be unmatched; the percentage of incarceration sentences is 12 percentage points lower in the final sample than in the full case data, and among incarceration sentences, the average sentence length is 6.10 months shorter. Regarding offense type distributions, violent and drug offenses are more prevalent in the final sample, and public order crimes are less prevalent. The distributions of offense severity classes are fairly similar, though Class C misdemeanors are substantially less likely to be in the final sample. The average number of prior adult convictions is smaller in the final sample (3.28 vs 4.04), and the average number of current convictions is slightly larger (1.68 vs 1.57). More cases in the final sample were charged with a violation of probation (15% vs 11%), and fewer were convicted of a mandatory minimum offense, but count reductions, severity reductions, and the distribution of conviction years in the final sample are all within 2% of the distribution in the full case data.

Regarding offender characteristics, the final sample is also more likely to be white (50% vs 40%) and less likely to be black (30% vs 36%) or Hispanic (20% vs 23%). Females are slightly more prevalent in the final sample, at 17% compared to 15% in the full case data. Finally, the final sample is younger than the full case data sample, with the 18-24 age group comprising 8% more of the final sample and other age groups each making up 2-4% less.

These differences provide evidence that more serious cases are less likely to have had a risk assessment conducted within one year of the conviction date and are less likely to be a part of the final sample. Not only are incarceration sentences less likely to end up in the final sample, but cases with *longer* incarceration sentences and longer criminal histories are less likely to do so. Cases with black, Hispanic, or male defendants are also less likely to be in the final sample, which is consistent with prior research indicating that minority and male offenders tend to be charged with more serious crimes and have more extensive criminal histories (Kramer & Steffensmeier, 1993; Kutateladze et al., 2015).

The researcher believes that this sample selection has the potential to influence the results in this study in two ways. First, sample selection may increase observed disparities in Connecticut's current sentencing scheme. Second, it may inflate the observed relationship between LSI-R scores and punishment in Connecticut's current sentencing scheme, which would in turn result in larger estimates of the relationship between risk and sentencing outcomes in the simulated evidence-based sentences. Each of these predictions will be addressed in turn, but both are drawn from the liberation hypothesis, which posits that judges feel constrained to issue harsh sentences when case seriousness is high; the extremity of legal characteristics like offense severity and criminal history crowd out the role that extralegal factors can play in determining punishment. In contrast, when case seriousness is low, judges feel liberated to incorporate extralegal considerations into their decision-making, and disparities are more likely to emerge (Kalven & Zeisel, 1966; Spohn & Cederblom, 1991). Though results are not entirely consistent, prior research tends to support this supposition (Koons-Witt et al., 2014; Hester & Hartman, 2017; Spohn, 2000; Rodriguez et al., 2006).

Because non-serious cases are more likely to be selected into the sample, it is first expected that disparities observed among the selected cases will be larger in magnitude than disparities that would be observed among the cases excluded from analysis. It follows from the liberation hypothesis that among those excluded cases, particularly the most serious of them, judges would feel more constrained by the severity of the offense and circumstances, by offenders' long criminal histories, and by other legal factors to issue equally harsh sentences to offenders regardless of their sociodemographic characteristics. If these cases could be matched to a complete risk assessment and included in the analyses, they would counteract some of whatever disparity is observed among less serious cases, and the observed disparities in punishment would be slightly smaller overall.

The tendency for more serious cases to be excluded from the sample also leads to an expectation that the relationship between LSI-R risk scores and punishment will be larger in the final sample than it would be among the excluded cases. Using the same liberation hypothesis logic, judges are likely to weight LSI-R risk scores less heavily in their decision-making in cases where the legal characteristics of the case already point to especially harsh punishment. In cases where offenders have a long string of prior felony convictions or committed Class A felonies, for example, the value of judges also knowing offenders' risk scores likely decreases considerably; judges are likely to issue similarly long incarceration sentences even if the offenders receive very low risk scores. Consequently, the relationship between risk scores and punishment should be smaller among the more serious excluded cases than among the less serious final sample cases. If the excluded cases could be matched to a risk assessment and included in analyses, the observed relationship between risk scores and sentences would be attenuated to a degree. This would be expected to result in a) a weaker relationship between risk and punishment and ultimately b) less punishment disparity due to sociodemographic differences in LSI-R scores in the simulated evidence-based sentencing scenario created in this study. Thus, selection into the final sample may have an influence on the results obtained in this study.

CHAPTER 8: RESULTS

This chapter presents the findings for all three phases of the analysis. It begins by exploring sociodemographic disparities in Connecticut's current sentencing scheme, observing both aggregate disparities and how much of those disparities can be explained by legal factors and case processing factors, as well as by risk factors captured in the LSI-R. The analyses then turn to evaluating disparities in the LSI-R and in potential evidence-based sentences. First, differences in mean LSI-R composite and domain scores are detailed. Then the results of the evidence-based sentence construction procedure are described, and disparities in the evidence-based sentences are assessed. The chapter concludes with a comparison of disparities in Connecticut's current scheme and disparities in potential evidence-based sentencing, to address whether or not evidence-based sentencing could alter patterns of social inequality in Connecticut if implemented.

Phase 1: Disparities in Connecticut's Current Sentencing Scheme

Aggregate Disparities. The first research question in this study asks to what extent there are aggregate sociodemographic disparities in sentencing outcomes in Connecticut's current sentencing scheme. Table 8.1 displays aggregate disparities in the decision to incarcerate and incarceration length, estimated using a negative binomial hurdle model. This model focuses on the primary independent variables and does not include any control variables to address the first research question.

The decision to incarcerate. Results indicate some noteworthy disparities in the decision to incarcerate. Interestingly, neither the coefficient for black offenders nor the coefficient for Hispanic offenders reaches statistical significance, though the odds of incarceration are 9% higher for black offenders than whites and 6% lower for Hispanic offenders than whites.[28] The effect for Asian offenders is very small and does not reach statistical significance either. The similarity between Asian and white offenders in this model may be due to similar behavior and similar treatment by criminal justice actors, but it may also be partly due to the relatively small number of Asian offenders in the sample. Gender exerts an especially strong effect on the odds of incarceration, as males have over twice the odds of being sentenced to incarceration.

There are also differences in the odds of incarceration based on SES components. Generally, more education decreases the likelihood of incarceration; compared to offenders who have not completed 10th grade, the odds of incarceration are 14% and 12% lower for offenders who have a high school degree and offenders who completed some college, respectively. Likewise, employment status exhibits a substantial effect on incarceration; compared to unemployed non-student offenders, part-time employed or student offenders have 50% lower odds of incarceration and full-time employed or student offenders have 53% lower odds. The effects of the third indicator of SES, financial situation, are less consistent. Reliance upon government assistance decreases the odds of incarceration by 10%, but each additional point in the 4-point financial problems scale increases the odds of incarceration by 17%.

Thus, there are several aggregate sociodemographic disparities in the decision to incarcerate, not all of which are consistent with previous literature on disparities in punishment. Disadvantages for males and low-SES individuals are well-documented in

---

[28] Because race and ethnicity are closely intertwined with SES (Howard, 1975; Steffensmeier, Ulmer, & Kramer, 1998), the researcher performed supplementary analyses to evaluate whether or not the inclusion of SES variables affected observed aggregate racial/ethnic disparities in punishment. When all SES indicators are excluded from the aggregate model, such that race, ethnicity, and gender are the only predictors of the decision to incarcerate, black offenders have statistically significant 11% higher odds and Hispanic offenders have 1% lower odds of incarceration than whites. Though these differences suggest that there is some overlap between race/ethnicity and SES, overall substantive conclusions about the relationships between race/ethnicity, SES, and punishment do not change.

the empirical literature (Baumer, 2013; Daly & Bordt, 1995; Spohn, 2000; Ulmer, 2012), and the findings that suggest a disadvantage for male, low-educated, unemployed, and financially unstable offenders in Connecticut sentencing align with that prior work. However, the observed relationship between race/ethnicity and incarceration runs counter to general conclusions from prior work. While much scholarship indicates a disadvantage for black and Hispanic offenders at sentencing (e.g. Mitchell, 2005; Spohn, 2000; Zatz, 2000), the effects for black and Hispanic offenders in this study do not reach statistical significance, and the Hispanic effect is in the opposite direction as expected. Additionally, given that some research finds disadvantages in punishment for low-income offenders (e.g. Mustard, 2001), and even specifically for offenders who rely on others for financial assistance (Wooldredge, 2010), it is surprising that in this sample, offenders who are reliant upon government assistance are less likely to be incarcerated than those who are not.[29]

Incarceration length. Results for the aggregate disparities sentence length model are also displayed in Table 8.1. Included in this part of the table are average marginal effects, which describe how many additional or fewer months an offender will receive, on average, if the offender is coded as a one (for dichotomous variables) or has one additional unit of a variable (for non-dichotomous variables). For ease of interpretation, it is the average marginal effects that are discussed in this text. When only the primary independent variables are included, aggregate disparities can be observed for race,

[29] Because this result is unexpected, the researcher conducted supplementary analyses to determine whether the observed effect for social assistance was impacted by the inclusion of a related measure, the financial problems scale. These two variables have a small-to-moderate correlation (.29), and the removal of the financial problems scale from the aggregate disparities incarceration model reduces the effect of receiving social assistance from an odds ratio of .90 to .94, which means that the direction of the social assistance effect is not due just to the inclusion of the financial problems scale.

ethnicity, gender, and every indicator of SES except for the financial problems scale. Unlike the model predicting the decision to incarcerate, in this model the effects for both black and Hispanic offenders are both positive and statistically significant. Black offenders receive, on average, 8.3 more months of incarceration than whites, while Hispanic offenders receive 6.25 more than whites.[30] The coefficient for Asian offenders is not even close to statistically significant due to a large standard error (again, likely due to the small number of Asian offenders represented in the data), though the point estimate indicates that Asian offenders also receive an average of 6.9 more months of incarceration than whites. The gender gap is again quite large; males receive over 7 more months of incarceration than females.

Disparities are present in each of the three indicators of SES as well. The coefficients for education follow a consistent pattern in which offenders with more education receive fewer months of incarceration. When compared to offenders who did not complete 10th grade, these differences range from a relatively small 2.0 months less for offenders who completed 10th grade up to 7.6 months less for offenders with some post-graduate education. Employed offenders also appear to have an advantage, as both part-time and full-time employed/student offenders receive approximately 7.4 months less incarceration time than unemployed/non-student offenders. Finally, disparities by financial situation follow a slightly different pattern than they did in the models predicting the decision to incarcerate. Reliance upon social assistance remains an advantage, reducing the incarceration sentence by an average of 5.8 months, but there is

[30] Supplementary analyses indicate that similarly to the decision to incarcerate, the observed effects of race/ethnicity on sentence lengths are marginally larger (9.21 and 7.90 months more for black and Hispanic offenders, respectively) when SES indicators are excluded from the model. This again suggests that there is some overlap between racial/ethnic and SES effects on incarceration length, but the general conclusions are similar regardless of whether SES indicators are included.

no discernible effect of additional points on the financial problems scale on sentence length.

These patterns of sociodemographic disparity are fairly similar to the patterns observed in disparities in the decision to incarcerate, and nearly all of them align with findings from prior empirical work (Bontrager et al., 2013; Franklin, 2017; Mitchell, 2005; Spohn, 2000; Wooldredge, 2010). With the exception of the effects for Asians and for offenders who are reliant upon social assistance, these results generally indicate disadvantages for racial and ethnic minority, male, and low-SES offenders. However, these aggregate disparities capture differences both in treatment by the criminal justice system and in offending behaviors; though discretion may be causing some of the disparity, black, Hispanic, male, and low-SES offenders may also commit more serious crimes, have more extensive criminal histories, or otherwise interact with the criminal justice system in ways that earn them harsher punishments. For this reason, it is important to distinguish between disparities due to court actor discretion and disparities due to legal and case processing characteristics.

Adding Legal and Case Processing Factors. Though Connecticut does not have sentencing guidelines to structure sentencing decisions, numerous case and case processing factors can (and in some instances are legally prescribed to) shape punishment outcomes. These factors may also serve as vehicles for sociodemographic disparities, so it is important to consider how much of observed disparities can be explained by these factors (RQ2). To address the second research question, in the next step of each stepwise regression, legal and case processing variables are added to the model. These additions are displayed in Table 8.2.

The decision to incarcerate. Adding legal and case processing factors does change several sociodemographic disparities in the decision to incarcerate. The effect of being black actually becomes statistically significant in the opposite direction. After accounting for legal and case processing factors, black offenders have 14% *lower* odds of being incarcerated than whites. Hispanic offenders remain less likely to be incarcerated than whites as well, and the magnitude of the difference actually grows large enough to achieve statistical significance. The effect for Asians is again not statistically significant. Regarding the gender effect, legal and case processing factors reduce the difference in odds of incarceration between male and female offenders by 26% (from 217% higher odds to 160% higher odds).

The effects of educational attainment are changed by the addition of legal and case processing factors as well. While the effect of completing 10^{th} grade remains insignificant, the significant advantage gained by earning a high school degree compared to completing less than 10^{th} grade is eliminated. The effect of completing some college changes direction, such that once legal and case processing factors are accounted for, college experience actually increases the odds of incarceration by 8% compared to completing less than 10^{th} grade. Interestingly, the education group with the highest odds of incarceration after legal and case processing factors are included is offenders with post-graduate education, whose odds of incarceration are now 49% higher than offenders who have completed less than 10^{th} grade. The effects of employment are reduced but not eliminated by the addition of legal and case processing variables, with reductions in magnitude of 22% (50% lower odds to 39% lower odds) for part-time employed/student offenders and 21% (53% lower odds to 42% lower odds) for full-time employed/student offenders. In terms of financial situation, the effect of reliance on social assistance increases in magnitude, with an odds ratio that shifts from .90 to .82. At the same time, legal and case processing factors do appear to explain some of the relationship between financial problems and incarceration; an extra point on the financial problems scale now increases the odds of incarceration by only 9%.

On the whole, the addition of legal and case processing characteristics to the incarceration model does little to "explain" aggregate sociodemographic disparities. The gender effect is notably reduced, as are the effects of employment and financial problems. This suggests that gender and some SES differences in punishment are partly due to differences in legal and case processing factors. But the effects of race, ethnicity, and education are changed in other less expected ways. The effect of being Hispanic is more than doubled in magnitude, and the effect of reliance of government assistance increases by 80%, while the effects of being black and having more education change direction entirely. In these cases, the findings suggest that legal and case processing factors are related to offender sociodemographic characteristics, but not in ways that simply intensify existing disparities.[31]

Many of the legal and case processing factors also have their own direct impact on the decision to incarcerate. Violent, drug, sex, and weapons crimes are all more likely to earn an incarceration sentence than public order crimes. Somewhat surprisingly,

[31] Supplemental analyses show that it is the offense severity classifications and/or criminal history measures that are primarily responsible for these shifts. The magnification of the Hispanic effect is due to offense severity; when offense severity categories are the only legal/case processing variables added to the model, Hispanics' odds of incarceration ratio drops all the way down to .82. The directional change in the effect for black offenders is driven by the addition of both criminal history and offense severity, as are most of the shifts in education effects. In support of these relationships, mean difference tests show that black, Hispanic, and less educated offenders are significantly more likely to commit more severe crimes (Kruskal-Wallis tests, $p<.001$ for all comparisons), while black, and less educated offenders have significantly more prior convictions (two-sample t test for black-white difference, ANOVA for education differences, $p<.001$ for both comparisons).

committing a violent crime appears to have a smaller effect than committing a drug, sex, or weapons crime. Property crimes do not appear to have a distinguishable effect on the likelihood of incarceration relative to public order crimes. The effects of offense severity generally align with reasonable expectations; the odds of incarceration are higher for all felony groups than for the reference category of Class A misdemeanors (the exception to this is the very small group of Class E felonies, the effect of which does not reach statistical significance). Moreover, as felony offense severity increases, the odds ratios get progressively larger, indicating larger differences in incarceration rates between more severe groups and the Class A misdemeanor reference group. On the flip side, the odds ratios for all other misdemeanor groups are less than 1, indicating that cases in those groups are all less likely to involve an incarceration sentence than cases with a Class A misdemeanor.

The two measures of criminal history appear to conjunctively have an effect on incarceration. Each additional prior conviction increases the odds of incarceration by 3%, while each additional point on the criminal history score increases the odds of incarceration by 19%.[32] Each additional current conviction count increases the odds of incarceration by 41%, and being charged with a violation of probation alongside a felony or misdemeanor increases the odds of incarceration by 83%. Cases involving a

[32] The researcher conducted supplemental analyses to consider the value of including the criminal history domain score as a second indicator of criminal history in the incarceration model. Prior adult convictions and domain scores correlate highly at .65, suggesting potential overlap in the predictive utility of the two criminal history measures. Indeed, when the criminal history domain score is excluded from the model, the odds ratio for prior convictions jumps from 1.03 to 1.10, more than a threefold increase in the effect of an additional prior conviction. When prior convictions are excluded from the model, the odds ratio for the criminal history domain score increases from 1.19 to 1.21, an 11% increase. However, Wald tests indicate that the criminal history domain score does significantly increase model fit when prior convictions are already in the model (χ^2 = 255.91, p<.001), and prior adult convictions significantly increase model fit when the criminal history domain score is already in the model (χ^2 = 6.95, p=.008). Therefore, the researcher decided to keep both measures of criminal history in the incarceration model, acknowledging that including both indicators dampens the observed effect of each on the likelihood of incarceration.

mandatory minimum also have odds of incarceration that are approximately 3.5 times greater. There appear to be some differences in incarceration across time too; cases with a verdict that occurred in FY 2009 or 2010 are more likely to involve an incarceration sentence than verdicts that occurred in FY 2008, though these effects do not reach statistical significance.

Finally, the indicators of plea bargaining show mixed effects; cases involving a reduction in counts or offense severity have 12% and 5% lower odds of incarceration, respectively, but cases in which a mandatory minimum offense is dropped have 18% higher odds. When the plea bargain involves a count or severity reduction, these results suggest that judges may be rewarding offenders for pleading guilty and avoiding the resources necessary for a trial. On the other hand, the finding that cases with a dropped mandatory minimum are more likely to result in incarceration than cases with no mandatory minimum may be an indication that even when an offender pleads guilty in order to avoid a mandatory minimum penalty, the severity of the initial charges still impacts judges' (and possibly prosecutors') assessments of blameworthiness and affects the final sentence issued. These two sets of findings are difficult to reconcile, but it may be that the indicators of plea bargaining used in this study are more influential for determining conviction offenses. Because conviction offenses are accounted for in the model, the effects of the plea bargaining indicators on punishment severity become less consistent.

Sentence length. The addition of legal and case processing factors to the sentence length model appears to substantially reduce the magnitude of sociodemographic disparities. The disadvantages for black and Hispanic offenders are more than cut in half, down from 8.3 extra months to 4.1 for blacks and from 6.3 extra months to 3.0 for Hispanics. The reduction in the gender gap is even larger. Once legal/case processing factors are added, the marginal effect for males drops from 7.2 months to 2.6, a 64% decrease. The effects for education groups also shrink by over 65%, and two of them (completion of 10th grade and post-graduate education) lose statistical significance. Similarly, the effects for part-time and full-time employment are reduced by 84% and 75% but remain significant, and the effect for reliance on government assistance is reduced by 74% but retains statistical significance. The effect for financial problems remains negligible.

Prior literature typically finds that large portions of sociodemographic disparity can be explained by offense and offense history characteristics (Spohn, 2000; Zatz, 2000), and these results are consistent with that trend. Contrary to findings from the decision to incarcerate model, the introduction of legal and case processing factors into the sentence length model does substantially reduce the magnitude of sociodemographic disparities in every instance where a statistically significant difference had been previously observed. Though legal and case processing factors do not account for *all* the observed disparities, they do provide support to the proposition that differences in those factors are responsible for some disparities in incarceration length.

Of course, the legal and case processing characteristics also have their own independent effects on sentence lengths. Violent, property, sex, and weapons crimes all tend to lead to more months of incarceration than public order crimes. Perhaps surprisingly, drug crimes tend to receive shorter sentences than public order crimes, by an average of 9.0 months. Similarly to the decision to incarcerate model, the relationship between offense severity classification and offense severity follows a predictable pattern in the incarceration length model. When compared to the Class A misdemeanor reference category, each increase in the severity level is associated with more months of incarceration. The most severe non-capital felony category, Class A murder, is associated with an average of 70 more months of incarceration than the Class A misdemeanor reference category, while the lowest misdemeanor category, Class D misdemeanor, is associated with 14.3 fewer months.[33]

Each additional prior conviction increases the incarceration length by only .5 months, but each additional point on the LSI-R Criminal History domain score increases the sentence by about .4 months.[34] Each additional conviction count increases the sentence length by 4.3 months. Committing a new crime while under probation supervision also increases sentence length; those charged with a probation violation receive sentences that are 4.2 months longer. The application of a mandatory minimum penalty increases the sentence by an average of 10.9 months as well. Once again, the indicators of plea bargaining show mixed effects. Cases in which one or more charges were dropped prior to conviction receive sentences that are over 5 months shorter. In contrast, cases involving a reduction in offense severity receive sentences that are over 6 months longer. Cases in which a mandatory minimum charge was dropped also receive

[33] The effect for unclassified misdemeanors, which was small and statistically non-significant in the decision to incarcerate models, is here very large: unclassified misdemeanors receive, on average, nearly 35 fewer months of incarceration. Because there are only 7 unclassified misdemeanors included in this sample, however, this point estimate should not be used to make inferences about the full population of unclassified misdemeanors.

[34] Here again the researcher used supplementary analysis to evaluate the utility of including both the Criminal History domain score and number of prior convictions in the model. Wald tests indicate that the criminal history domain score does significantly increase length model fit when prior convictions are already in the model ($\chi^2 = 33.96$, $p<.001$), and prior adult convictions significantly increase length model fit when the criminal history domain score is already in the model ($\chi^2 = 118.13$, $p<.001$). Therefore, both indicators of criminal history are kept in this model, with the understanding that the two indicators somewhat overlap in the sentence length variation they explain.

sentences that are over 4.5 months longer. Again, these contrasting findings may be a signal that the plea bargaining indicators used in this study are more important for determining conviction offense severity than for determining punishment severity, and once conviction offense severity is controlled for, the effects of the plea bargaining indicators become less meaningful. Sentence lengths also appear to get slightly longer as time moves forward; convictions in FY 2009 and 2010 earn sentences that are about .5 and 1.1 months longer than convictions in 2008, though only the effect for FY 2010 reaches statistical significance.

Adding Risk Factors. Especially in a non-guidelines state like Connecticut, where judges are afforded substantial discretion at sentencing, sociodemographic disparities may emerge indirectly through factors other than legal and case processing characteristics. The third research question asks whether differences in extralegal factors that have been empirically linked to risk of recidivism are responsible for some of the observed disparities in punishment. To examine this possibility, in the final step of the stepwise regression analyses, age, marital status, repeat convictions, and seven risk domain scores from the LSI-R are added to the models. These full models are displayed in Table 8.3.[35]

The decision to incarcerate. Adding risk factors into the incarceration model has inconsistent effects on sociodemographic disparities. The racial and ethnic effects on incarceration grow larger, such that black offenders have 20% lower odds and Hispanic

[35] As a supplementary analysis, the researcher also estimated models in which composite LSI-R risk scores are used as a risk factor in place of the seven individual risk domains scores [results not shown]. When the composite score was included instead of the individual domain scores, observed sociodemographic disparities were all in the same direction but tended to be marginally larger, suggesting that the group of individual risk domains contribute more to the explanation of social inequality in punishment than the composite score alone. The composite score also had its own significant effect on punishment, with each additional point increasing the odds of incarceration by 1% and the expected sentence length by .21 months.

offenders have 17% lower odds of incarceration compared to whites after accounting for legal, case processing, and risk factors. The effect for Asians remains insignificant. Adding risk factors reduces the gender gap by about 15%, changing the odds ratio from 1.60 to 1.51, but the effect of gender on incarceration remains large and significant.

Turning to SES components, risk factors only seem to impact the effects of financial situation. For education, the addition of the risk factors does little to the gap in incarceration between low-educated and high-educated offenders; the effects of completing 10th grade and having a high school degree remain non-significant, and the effects of completing some college or post-graduate education remain positive and approximately the same magnitude. Likewise, adding risk factors to the model narrows the employment gap only slightly; the effects of part-time and full-time employment/student status on incarceration decrease by rather small margins. In contrast, though the effect of reliance upon social assistance barely changes at all, the addition of the other risk factors reduces the effect of an additional point on the financial problems scale by 56% and renders it statistically non-significant.

Overall, the effects of risk factors on sociodemographic disparities in the decision to incarcerate are quite varied. Some conform to expectations. For example, accounting for risk factors that previous research indicates are more prevalent for racial or ethnic minority, males, young, and low-SES offenders (see Manchak et al., 2009; Monahan, Skeem, & Lowenkamp, 2017; Skeem & Lowenkamp, 2016; Skeem, Monahan, & Lowenkamp, 2016) reduces the magnitude of the disadvantage that male and financially unstable offenders have at sentencing. This suggests that some of the sentencing disadvantage is due to those groups being identified or perceived as higher-risk. Though the direction of the race effect is unexpected, the direction of the *change* in the race effect after accounting for risk factors is also consistent with expectations. The race effect becomes more negative, an indication that even though black offenders in this sample may receive an advantage at sentencing, they may still be perceived as having a higher risk of reoffending. At the same time, the impacts of education, employment, and receiving social assistance on incarceration are largely unchanged by the addition of the risk factors. It may be that because judges face such immense time and resources constraints at sentencing, easily observable factors that signal lower SES become especially salient as signals for blameworthiness, dangerousness, and even ability to withstand incarceration, leading to a large role for those factors independent of informal risk evaluation.

Adding risk factors to the model does little to change the effect of legal and case processing characteristics on incarceration either. Though many of the odds ratios shift, the changes are quite small and do not impact substantive conclusions about legal/case processing characteristics' influence on the decision to incarcerate. However, many of the added risk factors do have their own unique effects on incarceration. Among age categories, only the comparison between the oldest and youngest offenders achieves statistical significance; offenders who are at least 45 years old have 12% higher odds of incarceration than offenders who are under 25. The effects for 25-34 year olds and 35-44 year olds are small and not statistically significant. Likewise, married offenders do not appear to have a significant advantage over single offenders in the decision to incarcerate. However, multiple appearances in the dataset do appear to have a significant impact on incarceration. Having a previous conviction during the study period increases the likelihood of incarceration by 14%, even after accounting for number of prior convictions.

There are also several noteworthy effects among the risk domains included in the LSI-R. Interestingly, the largest observed per-point effect lies in the Accommodation domain, which captures both residential stability and neighborhood crime rates. An additional point in this scale increases the odds of incarceration by a sizeable 14%. Several of the domain scores, including Leisure/Recreation, Companions, Alcohol/Drug, and Attitudes/Orientations, have more modest interval impacts on the likelihood of incarceration and do not reach statistical significance, with changes in odds between -1% and 4%.[36] Finally, the Family/Marital domain score and Emotional/Personal domain score actually have negative significant impacts on sentencing; each additional point on the Family/Marital scale decreases the odds of incarceration by 3%, and each additional point on the Emotional/Personal domain score decreases the odds by 5%. These negative effects are somewhat surprising, but the effect for the Emotional/Personal domain may due to the fact that the indicators in that domain all relate to mental illness diagnoses and mental health treatment. Though poor mental health may indeed predict recidivism, judges may also see a diagnosis as a factor that mitigates the severity of the offense or the blameworthiness of the offender. Moreover, judges may be more inclined to allow community supervision for offenders who are already receiving mental health treatment, so that an incarceration stint will not interfere with their progress.

---

[36] Note however, that because domains contain different numbers of items, the maximum effect of one domain score on incarceration varies substantially across domains even when odds ratios are similar. Setting statistical significance aside, an odds ratio of 1.01 for the Alcohol/Drug domain means that each additional point increases the odds of incarceration by only 1%, for example, but because there are nine items in the domain, the difference in incarceration likelihoods between offenders at the very top and bottom of the scale is larger than it is for offenders at the top and bottom of the Companions scale, which is associated with 1% lower incarceration odds per point but only has five items.

***Sentence length*.** The addition of risk factors also has mixed effects on disparities in sentence lengths. Age, marital status, repeat convictions in the dataset, and LSI-R risk domains appear to account for about 19% of the remaining disparity between black and white offenders, reducing the extra months of incarceration for blacks from 4.1 to 3.3. The coefficient remains statistically significant. For Hispanics, the addition reduces the incarceration length gap from 3.0 to 2.5 months, another modest 18% reduction that does not result in a loss of statistical significance. Surprisingly, risk factors do almost nothing to account for the remaining gender gap in incarceration; the coefficient for male offenders decreases by only .06 months, or 2%, with the addition of those factors. After accounting for both legal/case processing and risk factors, the gender gap remains.

In terms of SES effects, the addition of risk factors noticeably attenuates the differences between more-educated and less-educated offenders as well as between unemployed non-student and full-time employee or student offenders. Specifically, including other risk factors reduces the advantage for offenders with a high school degree by 32% (from 1.54 to 1.05 fewer months), the advantage for offenders with a college degree by 48% (from 1.97 to 1.03 fewer months), and the advantage for full-time employed/student offenders by 24% (from 1.87 to 1.42 fewer months). There does not appear to be any change in the effect of reliance on government assistance, which remains statistically significant, or the effect of an additional point on the financial problems scale, which remains minimal and non-significant.

Thus, differences in recidivism risk seem to account for some residual disparity in incarceration length, but the risk factors captured in this analysis fail to account for all of it. Risk can explain more racial/ethnic disparity than gender disparity in sentence length, and the effects of educational attainment, employment status, and financial problems are attenuated but not eliminated by the addition of the risk factors. Even with all risk factors included in the model, race, gender, education, employment, and reliance upon social assistance continue to have influences on incarceration length. These findings are consistent with theoretical expectations, which predict that sociodemographic disparities can be observed even after accounting for legal and case processing factors because stereotypes that disadvantage certain groups are used to make crime-related attributions about offenders at the time of sentencing (Albonetti, 1991; Steffensmeier et al., 1996).

The effects of legal and case processing characteristics change relatively little once risk factors are included in the model. Turning now to the direct effects of risk factors on length of incarceration, age again appears to have very little effect on sentence length. Only one of the three comparisons reaches statistical significance; offenders who are 45 years or older receive an average of about 2.4 fewer months of incarceration than offenders who are under 25. Married offenders receive sentences that are .7 months longer, on average, but the effect does not reach statistical significance. Having been previously convicted during the study period does not significantly change the average sentence length when criminal history is already accounted for.

Among the LSI-R risk domains, the effects on incarceration length are all quite modest, and only some of them achieve statistical significance. The Companions domain score, which has virtually no effect on the decision to incarcerate, has the largest per-point effect on incarceration length; each additional point in the Companions scale increases the average sentence length by 1.1 months. The other significant positive effects are in the Attitudes/Orientations domain, where each additional point is associated with .50 months of extra incarceration time, and in the Emotional/Personal domains, where an additional point increases the sentence length by .23 months. The Accommodations domain also appears to have a comparable effect, at .34 months extra per point, though this effect does not reach statistical significance. In two domains, higher scores are actually associated with lower sentences; in the Leisure/Recreation domain, an additional point actually reduces the sentence length by .57 months, and in the Alcohol/Drug domain, an additional point is associated with a .19 month decrease. Finally, the effect for Family/Marital is also slight and does not reach statistical significance.

Summary of Disparities in Connecticut's Current Sentencing Scheme. In sum, patterns of disparity vary between the decision to incarcerate and sentence length. In the decision to incarcerate, racial and ethnic disparities are not observed in the aggregate, but accounting for legal and case processing factors and then risk factors actually uncovers an advantage for black and Hispanic offenders, which runs counter to expectations based on prior work (Baumer, 2013; Mitchell, 2005; Spohn, 2000). In contrast, black and Hispanic offenders are consistently disadvantaged in the sentence length model; even though the addition of legal, case processing, and risk factors lessen the sentence length gap between these groups and whites, the differences do not disappear entirely. The pattern is more consistent between the two punishment outcomes for gender, such that males are more likely to be incarcerated and receive longer incarceration sentences. In both models, legal and case processing factors account for some of this disparity, and risk factors account for very little, but a gender gap remains.

Differences in SES disparities between the decision to incarcerate and sentence length depend upon which indicator of SES is considered. When education is the focus, the pattern differs between outcomes. Disparities in the decision to incarcerate according to educational attainment follow an unexpected pattern; in the aggregate, higher levels of education are associated with lower odds of incarceration, a pattern that is consistent with prior literature (Albonetti, 1997; Franklin, 2017), but accounting for legal/case processing factors and risk factors exposes a disadvantage for the most highly educated offenders. On the other hand, there are aggregate disparities in sentence lengths that also favor higher-educated offenders, but legal/case processing and risk factors appear to explain some of those disparities rather than reverse their direction.

Patterns of disparity are likewise different between the two punishment outcomes for the other indicators of SES. In both models, greater levels of employment serve as an advantage in the aggregate; being a part-time or full-time employee/student decreases the odds of incarceration and sentence lengths. The addition of legal, case processing, and risk factors account for some of these disparities, but employed offenders retain an advantage even in the fully specified models. These trends are generally consistent with prior literature, which often suggests a persistent disadvantage for unemployed offenders (Chiricos & Bales, 1991; Myers, 1987; Wooldredge, 2010). Counter to expectations (e.g. Wooldredge, 2010), reliance upon social assistance is associated with an aggregate advantage in both the decision to incarcerate and sentence length, but the addition of legal/case processing factors and risk factors increases the magnitude of the effect on the decision to incarcerate and decreases the effect on sentence length. Finally, having a more satisfactory financial situation is likewise only an advantage in the decision to incarcerate; additional points on the financial problems scale increases the likelihood of incarceration but not sentence lengths, though the effect for the incarceration decision is rendered non-significant by the addition of legal/case processing and risk factors.

These results provide evidence that there are important sociodemographic disparities in Connecticut's current sentencing system. For the decision to incarcerate, aggregate disparities are noticeable by gender, education level, employment status, and financial situation. For incarceration length, there are aggregate disparities in every sociodemographic characteristic considered: race, ethnicity, gender, education level, employment status, and financial situation. These findings confirm that there are pre-existing disparities in punishment that must be taken into account in this evaluation of the impact of evidence-based sentencing on social inequality. With that in mind, the analysis now turns toward disparities in LSI-R scores and in simulated evidence-based sentences.

Phase 2: Disparities in LSI-R Scores and Evidence-Based Sentences

Disparities in LSI-R Scores. Scholars argue that evidence-based sentencing generates sociodemographic disparities in sentencing because the correlates of recidivism captured in risk assessment are also correlated with defendants' demographic characteristics (Hannah-Moffat, 2005, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b). To evaluate this proposition, it is first necessary to determine whether the underlying assumption that there are disparities in risk assessment is accurate. The fourth research question in this study asks whether there are sociodemographic disparities in LSI-R composite and domain scores.

Aggregate disparities. Table 8.4 first shows differences in average LSI-R composite scores by each racial/ethnic, gender, and SES comparison. It is clear that there are substantial aggregate disparities in composite scores. The largest racial difference is in the Asian-white comparison: Asian defendants receive average composite scores that are over five points lower than white defendants. Blacks receive an average score that 1.17 points higher than whites, and for Hispanics, the increase in risk is smaller, at .65 additional points. There is virtually no difference in average composite scores between male and female offenders. Because race, ethnicity, and gender are not directly captured in the LSI-R, differences between racial/ethnic and gender groups are attributable to racial/ethnic and gender differences in items that are captured in the LSI-R assessment.

Fully interpreting disparities among components of SES requires an understanding that some of these variables are drawn directly from the LSI-R risk assessment itself. Specifically, an offender earns one additional point each on the composite score if he/she is 1) currently unemployed, 2) has not completed grade 10, 3) has not completed grade 12, 4) has an unsatisfactory financial situation, or 5) is reliant upon social assistance. Therefore, the base expectation for disparities based on educational attainment is that compared to offenders who have not completed 10^{th} grade, offenders who have completed 10^{th} grade will receive one less point on the LSI-R, and offenders who have completed high school, some college, or post-grad education will receive two less points. Regarding employment, the base expectation for disparities is that compared to unemployed non-student offenders, offenders who have part-time or full-time employment will receive an average of one less point on the LSI-R. Reliance upon social assistance is expected to show a one-point average increase in risk scores as well. Finally, because the 4-point financial problems scale is collapsed into a binary indicator of financial satisfaction for the purposes of LSI-R scoring ("very satisfactory" and “relatively satisfactory” finances are given zero points, “very unsatisfactory” and “relatively unsatisfactory” are given one point), it is expected that compared to offenders with very satisfactory finances, offenders with relatively unsatisfactory or very unsatisfactory finances will have an average of one additional point on the LSI-R. Offenders with relatively satisfactory finances are expected to have the same average score as offenders with very satisfactory finances. Disparities beyond these expectations are attributable to covariance between race/gender/education/employment/finances and other items in the LSI-R.

A clear and consistent pattern in composite scores emerges across all indicators of SES; low-SES individuals receive higher average risk scores regardless of which SES indicator is considered. For educational attainment, higher levels of education lead to lower risk scores. The difference between offenders who have and have not completed 10th grade is approximately what is expected (completing 10th grade yields 1.1 fewer points, compared to the expected one point). However, the disparity compared to the reference group grows progressively larger as educational attainment increases, and the differences are not attributable to the two LSI-R questions that directly capture the completion of 10th and 12th grade. Completing high school nets the average offender over four fewer risk points than the reference group, when the expectation based on directly-captured education items is that the difference will be only two. Offenders with college and post-graduate education experience earn 6.36 and 11.73 fewer points, respectively, when the expectation is still only two.

There is a similar trend for employment. Compared to unemployed, non-student offenders, part-timers earn an average of 3.6 fewer risk points, when the base expectation is one less point. The difference for full-time employed/student offenders is even larger, as they earn 10.8 fewer points on average, and the base expectation is still only one less point. Reliance on social assistance is associated with 6.3 additional points, compared to the expected one additional point. Lastly, the financial problems scale exhibits the same pattern as educational attainment and employment; the difference in average risk scores gets progressively larger as the comparison groups become less similar. Compared to offenders with very satisfactory finances, offenders who have relatively satisfactory finances receive 3.8 fewer risk points, on average. This difference is not directly captured in the LSI-R, so the base expectation is that there would be no difference in average scores between these two groups. A one-point difference is expected for the relatively unsatisfactory and very unsatisfactory finances groups compared to the reference group, however, and both comparisons far surpass this expectation: offenders with relatively unsatisfactory finances earn 10.5 fewer points, and offenders with very unsatisfactory finances earn an average of nearly 14 fewer points.

These observations indicate that there are indeed disparities in LSI-R composite scores. Though race and ethnicity are not directly captured in the LSI-R, black and Hispanic offenders receive higher average risk scores than whites, while Asian offenders receive substantially lower scores. This is consistent with prior work, which often finds a risk score gap between racial/ethnic minorities and whites (Manchak et al., 2009; Monahan, Skeem, & Lowenkamp, 2017). However, the magnitude of the disadvantage for black and Hispanic offenders is quite modest, and it provides at most tepid support for critics of risk assessments and evidence-based sentencing, who often focus their concerns on disparate impacts for racial and ethnic minority offenders (e.g. Harcourt, 2007; Starr, 2014a,b). Interestingly, the difference in composite scores between male and female offenders is negligible, which is perhaps only slightly surprising, as prior research looking at the gender gap in risk scores yields mixed conclusions (Andrews & Bonta, 2003; Lowenkamp et al., 2001; Mihailides et al., 2005). The trends for educational attainment, employment status, and financial situation all support the conclusion that low-SES offenders receive higher LSI-R scores. Moreover, these differences are not attributable entirely to the individual LSI-R items that directly measure them; low-educated offenders, unemployed offenders, and offenders with less satisfactory financial situations must receive higher average scores in at least some of the other LSI-R domains. To determine which specific domains contribute to each of the observed disparities, differences in domain scores are evaluated next.

Table 8.4 also shows differences by race, ethnicity, gender, and indicators of SES in the 10 LSI-R domain scores. The largest domain score differences between black and white offenders exist in the Criminal History and Education/Employment domains, which disadvantage black offenders, and in the Alcohol/Drug and Emotional/Personal domains, which actually appear to counteract that disadvantage for blacks. Black offenders also have small disadvantages, ranging from .04 to .37 additional points, in the other six domains, which explains why the overall composite score indicates a relatively modest disadvantage for black offenders. For Hispanics, the largest disadvantage is centered in the Education/Employment domain and is nearly double the magnitude of the composite difference. The next two largest domain differences are in the Alcohol/Drug and Emotional/Personal domains, where Hispanics have significantly lower scores than whites. Much like in the black-white comparison, the remaining domain scores are all small and favor whites, which leads to a relatively small composite disadvantage for Hispanics. Comparisons between Asian and white offenders are fairly consistent across domains; Asian offenders receive lower domain scores in seven of the ten domains, and among two of the remaining domains, the difference does not even reach statistical significance. The largest differences are in the Criminal History and Alcohol/Drug domains, where Asian offenders average 1.26 and 1.38 points lower than whites.

Among the racial/ethnic domain comparisons, then, patterns for black and Hispanic offenders are very similar; the Education/Employment domain contributes the most to composite disadvantages for black and Hispanic offenders, and lower scores on the Alcohol/Drug and Emotional/Personal domains appear to mitigate much of what would have otherwise been large disadvantages in composite risk scores. These findings are fairly consistent with what little prior work there is; for example, Chenane and colleagues (2015) found that both black and Hispanic inmates had higher Education/Employment scores and lower Alcohol/Drug and Emotional/Personal scores than white inmates. Asian offenders, on the other hand, accumulate substantial advantages in several domains that are not counteracted in any meaningful way; differences in the Criminal History and Alcohol/Drug domains alone given Asian offenders an advantage of over 2.6 points, and other domains differences only add to that gap.

Turning to gender differences in domain scores, the two most prominent score differences counteract each other. They are in the Criminal History domain, where men score .87 points higher, and the Emotional/Personal domain, where men score .84 points lower. Men also score higher in the Education/Employment, Leisure/Recreation, Companions, Alcohol/Drug, and Attitudes/Orientations domains, while women score higher on the Financial, Family/Marital, and Accommodations domains, though the differences are not as stark. These results are also consistent with prior work; studies often find that men score higher on Criminal History, while women score higher on the Financial and Emotional/Personal scales (e.g. Manchak et al., 2009; Mihailides et al., 2005). The nearly non-existent gender difference in composite scores, then, masks several more noticeable domain disparities that suggest different risk factors are more prominent for male or female offenders.

As was the case with the composite score trends, the trends in domain scores for each SES indicator are remarkably consistent: low-SES offenders receive higher average domain scores. This is true in nearly every domain for every pairwise comparison, regardless of whether education, employment, reliance on social assistance, or financial situation is used as the indicator of SES.[37] As could reasonably be expected, the largest differences for several SES indicators are in the Education/Employment category, but while some of the differences can be explained by items that directly measure the education or employment status of the offender, the differences exceed the contribution of those particular items in most instances (e.g. two items in the Education/Employment domain measure educational attainment, but the differences between education groups are more than two points in all applicable comparisons). Other notable sources of composite score disparity are the Finances domain, which is subject to the same

[37] The most notable exception to this trend is in Alcohol/Drug domain scores compared across education groups; offenders with a 10th grade education, a high school degree, or some college experience receive higher risk scores than offenders with less than a 10th grade education, while offenders with a post-grad education receive lower scores.

individual item considerations as Education/Employment, and the Criminal History domain.

The findings for SES indicators show that low-SES offenders are at a substantial disadvantage in the LSI-R risk evaluation process. Not only are educational attainment, employment status, reliance on social assistance, and financial situation captured directly in the LSI-R, they additionally influence risk scores indirectly by increasing the likelihood of higher scores in all other domains. This is not an entirely new revelation. Scholars have argued that traditional indicators of recidivism risk are often proxies for social class (e.g. Silver & Miller, 2002). This practice is grounded in data-supported offending patterns; there is ample research showing that low SES is predictive of future criminality (see Gendreau et al, 1996 for meta-analytic evidence). That said, the marked consistency of these domain-specific findings within an assessment tool like the LSI-R, which identifies a seemingly diverse range of criminogenic factors, provides strong evidence that risk evaluation can easily become an exercise in identifying low-class offenders.

Disparities accounting for other sociodemographic factors. It is possible that these aggregate sociodemographic disparities may overlap statistically with each other; for example, there may be substantial shared variance in LSI-R scores among the separate indicators of SES. To account for this possibility and identify the unique relationship between each sociodemographic characteristic and the LSI-R, multivariate regression analyses assess the joint impact of race, ethnicity, gender, and SES on risk scores. Table 8.5 shows the results of these analyses for composite and domain scores on the LSI-R.

First, an ordinary least squares regression shows that significant (though sometimes modest) sociodemographic disparities exist for all characteristics examined. After controlling for other sociodemographics, black and Hispanic offenders receive slightly lower risk scores than whites: .32 and 1.58 points lower, respectively. For Asians, the advantage is slightly larger, at nearly three points. Consistent with the descriptive statistics, the gender gap remains relatively small in the regression model, with males earning only 1.2 points more than females. The indicators of SES, on the other hand, again show a consistent disadvantage for low-SES offenders that grows progressively larger as education, employment, and financial stability lessen. Though the differences here are not as large as the aggregate disparities, they are still substantial. The education difference rises to well over eight points between the highest- and lowest-educated offenders, for employment the gap is nearly six points between the full-time employed and unemployed offenders, social assistance is associated with over three additional points, and the top and bottom of the financial situation scale differ by 8.7 points. These findings reinforce the conclusions drawn from the consideration of aggregate disparities in composite scores; while composite scores are similar among racial, ethnic, and gender groups, indicators of SES maintain strong and consistent associations.

Ordered logistic regressions are also used in Table 8.5 to evaluate each domain from the LSI-R in turn. Unlike the ordinary least squares model, these models focus on the probability of membership in higher or lower categories rather than on risk point differences, so they are not directly comparable to the aggregate difference estimates detailed in Table 8.4. However, these models still provide valuable additional information about the impact of various sociodemographic characteristics on LSI-R domain scores.

Specifically, most of the relationships between race, ethnicity, gender, and SES categories and LSI-R scores observed in the multivariate ordered logistic regression models are in the same direction as they were observed in the aggregate statistics. To illustrate, the logistic model predicting Criminal History scores still shows that black, male, less educated, unemployed, assistance-reliant, and financially unstable offenders tend to receive higher scores. Interestingly, after controlling for gender and SES, Hispanic offenders are less likely than whites to be in a higher Criminal History category, a finding that stands in direct contrast to the aggregate descriptive statistics. A few other inconsistencies like this between the aggregate and multivariate analyses can be identified. For example, black and Hispanic offenders are less likely to receive higher Family/Marital scores only in the multivariate regression, and some of the relationships for Asian, low-educated, and assistance-reliant offenders that are significant in the aggregate statistics do not reach statistical significance in multivariate regressions. Generally speaking, though, the results from the multivariate analysis reinforce the conclusions drawn from the aggregate descriptive analysis: while domain scores do not consistently favor any particular racial, ethnic, or gender group, low-educated, unemployed, assistance-reliant, and financially unstable offenders receive higher scores across nearly all LSI-R domains.

Summary of disparities in the LSI-R. Overall, there are composite and domain disparities in LSI-R scores. The composite differences are modestly disadvantageous for black and Hispanic offenders, much more so for low-SES offenders. Asian offenders receive substantially lower scores than any other racial/ethnic group, and there is not much of any difference between male and female offenders' composite scores. While the risk domain patterns for race/ethnicity and gender are a mixed bag—some domains yield higher scores for black, Hispanic, Asian, and male offenders, while other domains yield lower scores—, the composite disadvantage for SES indicators is driven by disadvantages in every domain. These composite disparities may translate into disparities in punishment in an evidence-based sentencing scheme. The next section of the analysis simulates evidence-based sentences that incorporate these LSI-R risk scores.

Simulating Evidence-Based Sentences. As discussed in Chapter 6, this study relies on the relationship between risk and punishment in Connecticut's current sentencing scheme to illustrate the potential effects of incorporating LSI-R scores into punishment decisions. The first step in the evidence-based simulation process is therefore to estimate this relationship. Table 8.6 shows results for both the in/out and sentence length decisions. In this negative binomial hurdle model, sentence length is regressed on legal/case processing factors and composite LSI-R scores. This model constrains all other potential determinants of sentencing, most notably race, ethnicity, gender, and SES, to 0.

The odds ratios and average marginal effects obtained in this model are very similar to those obtained in Tables 8.2 and 8.3, when the sociodemographic variables of interest were also included as predictors. In this model, a relationship between composite LSI-R scores and punishment can also be observed. Each additional point on the LSI-R is associated with 2% higher odds of incarceration and with .18 months of additional incarceration time, on average. These marginal increases represent the relationship between aggregate risk and punishment in Connecticut's current system. Though they seem trivial at first glance, it is worth remembering that in this sample, LSI-R scores range from 0 to 51 (out of a possible 54); all else equal, the highest-risk offenders who scored a 51 are expected to have over twice the odds of incarceration and over nine additional months of incarceration length than the lowest-risk offenders who scored a 0.[38]

From this model, the researcher first uses a post-estimation procedure to predict the likelihood of incarceration for each offender.[39] For each offender in the sample, this procedure yields an expected probability between 0 and 1 that the offender would be incarcerated, given his/her legal/case processing characteristics and LSI-R composite score. These predicted probabilities can be interpreted as each offender's likelihood of incarceration in a system where only legal/case processing characteristics and LSI-R composite scores are allowed to determine punishment. These predicted probabilities are therefore treated as the predicted probabilities of incarceration for each offender in a simplified evidence-based sentencing scheme.

Using these predicted probabilities, the researcher assigns each offender either an evidence-based incarceration or non-incarceration sentence. As discussed in Chapter 6, a threshold rule is used to make these assignments. The researcher begins with a .5 threshold rule: if the offender has a probability of at least .5 of being incarcerated based on legal/case processing factors and LSI-R scores alone, the offender is assigned

[38] Supplementary analyses suggest that the relationship between LSI-R composite scores and punishment is larger when legal and case processing variables are excluded from the model. When composite scores are the only predictor of punishment, each additional point on the LSI-R composite score is associated with a 6% increase in the odds of incarceration and .31 additional months of incarceration. This translates into a threefold difference in the odds of incarceration and over 15 months of incarceration on average between offenders at the top and bottom of the LSI-R distribution. This provides some evidence that legal and case processing characteristics typically captured in official court records do partially tap into elements of offenders' risk of recidivism.

[39] The hurdle model used to model punishment in this study is not supported by post-estimation prediction commands in Stata. Therefore, in order to generate these predictions, the researcher separately estimated a logistic regression that predicted the in/out decision using the same independent variables. As a consistency check, the coefficients from the separate logistic regression were compared to the coefficients obtained from the logistic component of the hurdle model; the coefficients were identical, indicating that the hurdle model produced the same estimates as the logistic regression did separately. Following the logistic regression, the Stata post-estimation command "predict" was entered in Stata. This command yielded predicted probabilities derived from the logistic regression.

incarceration. If not, the offender is assigned non-incarceration. However, this decision rule is expected to yield an artificially low proportion of evidence-based incarceration sentences (see Footnote 25), which may in turn impact punishment disparities observed in the evidence-based system. To consider the sensitivity of this study's findings to the .5 threshold, thresholds of .4 and .6 are treated as alternative specifications as well.

The next step in the evidence-based sentencing simulation process requires assigning a sentence length to each evidence-based incarceration sentence. Returning to the model detailed in Table 8.6, the researcher uses a second post-estimation procedure to predict an expected incarceration length to each offender.[40] This procedure yields an expected sentence length for each offender if incarcerated, given his/her legal/case processing characteristics and LSI-R composite score. Offenders who were assigned an evidence-based incarceration sentence using the threshold decision rule are then assigned their evidence-based incarceration lengths. Offenders who were assigned an evidence-based non-incarceration sentence are assigned an evidence-based incarceration length of 0.

In the final step of the simulation process, evidence-based sentences issued in cases involving a mandatory minimum conviction are adjusted to meet the mandatory minimum penalty. For cases in which the mandatory minimum penalty exceeds the

[40] Again, because the hurdle model is not supported by post-estimation prediction commands, the author separately estimated a zero-truncated negative binomial model in which sentence length was regressed on all predictors for the sample of offenders who were issued a non-zero incarceration sentence. The Stata "predict" command was then used to generate the expected sentence length based on legal/case processing characteristics and LSI-R composite scores for each offender in the full sample.

evidence-based sentence issued, the mandatory minimum penalty becomes the new evidence-based sentence.[41]

Table 8.7 shows key statistics for the resultant simulated evidence-based sentencing distributions. As expected, the use of the .5 threshold leads to a substantially smaller group of evidence-based incarceration sentences; only 23% of offenders are issued an incarceration sentence in this scenario, compared to the 34% of offenders in Connecticut who were actually incarcerated. The .4 threshold nets the most similar proportion of incarcerated offenders, at 33%, and the .6 threshold produces an extremely low proportion, at just 15% incarcerated.

Looking at average sentence lengths among evidence-based incarcerated offenders, the average length in each threshold scenario is higher than the average sentence length actually received. As the probability threshold for evidence-based incarceration increases, so does the average sentence length of the offenders who are incarcerated in the evidence-based system. Even in the lowest threshold scenario, where offenders are assigned incarceration if their probability of evidence-based incarceration is at least .4, the average sentence length is over five months longer than the actual average sentence lengths received. This may be an indication that in Connecticut, many offenders with a low probability of incarceration based on their legal/case processing factors and risk of recidivism are sentenced to short stints in jail or prison rather than being assigned alternative sanctions. In fact, of the 22,515 offenders in the sample who were issued an actual incarceration sentence, 9,403 (42%) were assigned a probability of evidence-based incarceration less than .4 [results not shown]. Furthermore, those 9,403 offenders were

[41] When the .5 threshold is used, 68 sentences are adjusted upward to meet a mandatory minimum. 75 sentences are adjusted when the .4 threshold is used, and 65 sentences are adjusted when the .6 threshold is used.

sentenced to an average of 6.68 months of actual incarceration, a figure well below the full group average of 19.22 months.[42] At the same time, it may also be an indication that Connecticut judges already divert many offenders with higher likelihoods of evidence-based incarceration based on their legal/case processing factors and risk levels away from incarceration, perhaps to rehabilitative programming.[43] Among the 44,469 offenders who were issued non-incarceration sentences, 2,822 (6%) were assigned a probability of evidence-based incarceration greater than .6, suggesting that at least a few of the offenders most likely to be incarcerated were diverted [results not shown].

Disparities in Simulated Evidence-Based Sentences

It is clear that the evidence-based sentencing simulation procedures generate simulated sentences that differ substantially from the sentences actually issued in Connecticut during the study period. However, it is unclear whether or not there are sociodemographic disparities in these simulated evidence-based sentences, as there are in Connecticut's current sentencing scheme. The fifth research question asks to what extent sociodemographic disparities can be observed in the simulated evidence-based sentencing scenario.

Table 8.8 displays aggregate disparities in the evidence-based decision to incarcerate and sentence length, drawn from a negative binomial hurdle model when the .4 incarceration threshold is used. No control variables are included in this model. The same model is also estimated under the .5 and .6 threshold conditions (results displayed

---

[42] At this point, however, it is important to remember that the study does not have a reliable way to identify sentences of time served in the dataset. It is therefore likely that some of the 9,403 cases recorded as receiving an incarceration sentence are cases where the offender was sentenced to a short period of time served, which may explain why the data suggests that many offenders with low probabilities of incarceration received short incarceration sentences.

[43] Though it would be both interesting and informative to consider what alternative sanctions/programming offenders with higher likelihoods of evidence-based incarceration who were issued non-incarceration sentences were assigned, this study is unable to capture those different outcomes.

in Tables 8.9 and 8.10). Though the coefficients for each variable are similar across conditions, there is a distinct pattern in which several of the disparities grow slightly larger as the incarceration probability threshold increases. To remain consistent with other facets of this study that are expected to lead to conservative estimates of evidence-based punishment disparities, the presentation of these results will focus on evidence-based disparities when the incarceration probability threshold of .4 is used. This condition produces the smallest observed disparities in punishment.[44]

Even in the .4 threshold condition, which has the smallest odds ratios, there are substantial differences in the likelihood of evidence-based incarceration based on race, gender, education, employment status, and financial situation. Black offenders have 38% higher odds of evidence-based incarceration compared to whites. Males have just shy of twice the odds of evidence-based incarceration compared to females. Trends in the odds ratios for education groups indicate that higher educational attainment leads to progressively lower odds of evidence-based incarceration, ranging from 22% to 52% lower, and only the comparison between offenders who completed 10th grade and offenders who did not fails to reach statistical (and undoubtedly practical) significance. Employment trends likewise show that part-time employment reduces the odds of evidence-based incarceration by 42%, and full-time employment reduces them by 49%. Reliance upon government assistance and additional points on the financial problems scale also both increase the likelihood of evidence-based incarceration by 19% and 26%, respectively. The one general domain in which some disparities do not emerge is

[44] This condition has the added benefit of yielding outcomes with the most similar distribution to Connecticut's current sentencing outcomes.

race/ethnicity: in the evidence-based sentencing models, Hispanic and Asian offenders do not have significantly higher or lower odds of incarceration than whites.

These results largely conform to expectations based on both prior literature and earlier findings from this study's own consideration of disparities in the LSI-R. As discussed in the results from Phase 1, prior work tends to find disadvantages for black, Hispanic, male, and low-SES offenders at sentencing, much of which can be explained by differences in legally relevant factors (Bontrager et al., 2013; Franklin, 2017; Mitchell, 2005; Spohn, 2000; Wooldredge, 2010). With the exception of Hispanic offenders, those same trends are observed here, where punishment variation is driven entirely by legal and case processing factors plus LSI-R composite scores. Black, Hispanic, and low-SES offenders in this study's sample also received higher LSI-R composite scores, which likely further contributed to disadvantages in evidence-based incarceration for those groups.

Turning now to disparities in evidence-based incarceration lengths, Table 8.8 also illustrates aggregate disparities resulting from the zero-truncated negative binomial component of the hurdle model under the .4 incarceration threshold. Again, a comparison between the three threshold conditions in Tables 8.8, 8.9, and 8.10 show that the .4 threshold model yields the smallest sociodemographic disparities (according to average marginal effects; negative binomial coefficients show a slightly less consistent trend across models), and the presentation of results focuses on disparities in this more conservative model. It should be noted this time, however, that the magnitude of the disparities grows substantially larger as the incarceration probability threshold increases.

Unlike the evidence-based decision to incarcerate model, the evidence-based length model reveals no major differences according to race or ethnicity. Black offenders receive, on average, only 2.6 additional months of incarceration than whites, while Hispanic offenders receive 1.7 more, but with large standard errors these differences do not reach statistical significance. Asians receive an average of 5.77 additional months compared to whites, but this difference is not statistically significant either. On the other hand, male offenders receive an average of 3.3 more months of incarceration than females, a difference which does achieve statistical significance.

There are SES disparities in evidence-based incarceration lengths as well. Among the education groups, more education is generally associated with less incarceration time. Compared to offenders who did not complete 10th grade, offenders who completed 10th grade would receive 2.6, offenders who obtained a high school degree or experienced some college would receive 3.5, and offenders who had post-graduate education would receive 5.2 fewer months of evidence-based incarceration, on average. Unemployment is also a disadvantage in evidence-based sentencing; compared to unemployed offenders, part-time and full-time employed/student offenders would receive an average of 5.2 and 4.6 fewer months of incarceration. In terms of financial situation, reliance on social assistance is associated with 5.4 fewer months of evidence-based incarceration, but points on the financial problems scale appear to have little relationship with it.

In this model, the biggest surprise is that there is relatively little evidence of racial or ethnic inequality in simulated evidence-based sentence lengths. None of the threshold conditions reveal significant disparities between black, Hispanic, Asian, and white offenders. Even though racial and ethnic disparities in the LSI-R were relatively small, legal/case processing factors and risk factors appear to generate a fair amount of racial and ethnic disparity in Connecticut's current system. So it would have been reasonable to expect that sentence lengths predicated entirely on those factors would have exhibited noteworthy disparities as well. However, these results indicate that racial (though not ethnic) inequality is centered more in the decision to incarcerate than in incarceration lengths in the evidence-based scenario. The findings for gender and SES effects in evidence-based sentence lengths are less surprising. Large portions of the disparities that disadvantage male and low-SES offenders in Connecticut's current sentencing scheme can be explained by differences in legal and case processing factors, and low-SES offenders receive substantially higher LSI-R composite scores (though the same cannot be said for males and the LSI-R) regardless of which SES indicator is considered. Gender and SES disparities in the simulated evidence-based sentence lengths were therefore expected.

The trends across the .4, .5, and .6 threshold models reveal interesting information about the distribution of evidence-based incarceration risk and sentence length among different sociodemographic groups. Because moving from a .4 incarceration threshold to a .6 incarceration threshold increases the magnitude of disparities in the decision to incarcerate, one can ascertain that black, male, and low-SES offenders are overrepresented more among the group of offenders with a risk of evidence-based incarceration over .6 than they are among offenders with a risk between .4 and .6 (who are assigned incarceration in the .4 model but not in the .6). In other words, incarceration disparities for black, male, and low-SES offenders are more noticeable at the upper end of the probability distribution, where offenders have very high likelihoods of being incarcerated in the simulated evidence-based scenario, than they are in the middle of the distribution where there is less certainty.

This inference can be demonstrated descriptively. Black offenders make up 33% of the offenders with a probability between .4 and .6, but 37% of offenders with a probability above .6. For white offenders, the opposite is true: 46% of the mid-range probability group is white, while only 42% of the high-probability group is. Male offenders likewise make up 86% of the mid-range probability group and 90% of the high-probability group. The same pattern extends to most of the lowest-SES group within each SES indicator. Offenders who are the least-educated, unemployed, and have the least satisfactory finances make up 17%, 72%, and 28% of the mid-range probability group but 19%, 77%, and 31% of the high-probability group. The opposite trends are observed for the highest-educated, full-time employed, and most financially satisfactory offenders. For those who are reliant on social assistance, the proportion is nearly the same between threshold groups (55% vs 56%).

Moreover, because moving from a .4 to a .6 threshold also increases the magnitude of disparities in the simulated evidence-based sentence lengths, one can use the same logic to deduce that the gaps between black and white, male and female, and low-SES and high-SES offenders in expected evidence-based sentence lengths are larger at the top of the spectrum than in the middle as well. Descriptive statistics again match this conclusion. The average difference between white and black offenders' expected sentence lengths is .21 months in the mid-probability group and 4.06 months in the high-probability group. For gender, the difference actually favors males in the mid-probability group by 1.23 months but favors women in the high-probability group by 7.67 months.

Similar trends exist within SES groups. These observations together suggest that it is among the more serious cases, where legal/case processing factors tend to be severe and/or risk scores tend to be high, that black, male, and low-SES offenders are the most disadvantaged.

By and large, the evidence-based sentences simulated in this study do contain their own set of sociodemographic disparities. In the evidence-based scenario's decision to incarcerate, black, male, and low-SES offenders would be at a substantial disadvantage. In terms of evidence-based sentence lengths, male and low-SES offenders would tend to receive a few additional months of incarceration. Absent context, these findings provide preliminary evidence that the implementation of evidence-based sentencing in Connecticut may lead to disparate outcomes that further disadvantages some traditionally marginalized groups, such as black and low-SES individuals. Does evidence-based sentencing have the potential to foster social inequality in Connecticut? Patterns of disparity in the evidence-based sentences simulated in this study provide some indirect evidence that points to yes. However, earlier findings in this study indicate that sociodemographic disparities exist under Connecticut's current sentencing scheme as well, which brings a different question to the forefront: does evidence-based sentencing have the potential to *worsen* existing social inequality in Connecticut? This more nuanced question requires consideration of disparities in simulated evidence-based sentences relative to disparities in Connecticut's current system.

Phase 3: Comparing Disparities in Actual and Simulated Sentences

The sixth and final research question in this study asks to what extent the implementation of an evidence-based sentencing system in Connecticut has the potential to make sociodemographic disparities in punishment larger or smaller. To get at this important question, the final phase of this study compares disparities observed in Connecticut's current sentencing system to disparities observed in the simulated evidence-based sentencing scenario. Table 8.11 displays the odds ratios observed in the actual and simulated evidence-based models for the decision to incarcerate, as well as arrows indicating whether each disparity is larger or smaller in the simulated evidence-based sentencing model.[45]

Based on these results, simulated evidence-based sentences do not exhibit disparities in the likelihood of incarceration that are consistently greater or smaller than disparities in actual sentences, but several significant differences do emerge. Black offenders are disadvantaged to a substantially greater extent in simulated evidence-based incarceration compared to whites, and while Hispanic offenders maintain a slight advantage in the observed decision to incarcerate, they have a slight disadvantage in the simulated evidence-based decision to incarcerate. The advantage for Asian offenders grows larger in the simulated sentences, but the effect never reaches statistical significance. On the other hand, the gender gap appears to be somewhat smaller in simulated evidence-based sentences, though males retain a significant disadvantage regardless.

The potential effects of evidence-based sentencing on SES disparities are likewise mixed. Disparities based on educational attainment are greater in the simulated evidence-based sentences; offenders who obtained a high school degree, completed at least some college, or had some post-graduate education experience all receive greater advantages in

[45] Because the .4 threshold leads to the most conservative estimates of disparities in evidence-based sentences and provides a sentence distribution most similar to the sentences observed in Connecticut, the .4 threshold model is used for all comparisons.

the evidence-based scenario. In contrast, the effects on punishment of employment status are statistically significantly smaller in the simulated incarceration decision, albeit by relatively modest margins. Evidence-based sentencing has mixed effects on the two financial indicators as well. Reliance upon social assistance is an advantage in the actual decision to incarcerate, but it appears as a noteworthy disadvantage in the evidence-based decision to incarcerate. Meanwhile, each additional point on the financial problems scale is associated with larger increases in the likelihood of incarceration in the evidence-based scenario.

Overall, for the decision to incarcerate, disparities according to race, ethnicity, educational attainment, and financial situation are greater in the evidence-based sentences simulated in this study; consistent with scholarly criticisms of evidence-based sentencing (Hannah-Moffat, 2005, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014a,b), black, Hispanic, less educated, and financially unstable offenders are disadvantaged more in the evidence-based decision to incarcerate than they were in Connecticut's existing sentencing scheme. The gender and employment gap, on the other hand, are slightly smaller in evidence-based sentences, though men and unemployed offenders are still significantly more likely to be incarcerated.[46]

Turning one more time back to sentence lengths, Table 8.12 displays the average marginal effects observed in actual and simulated evidence-based sentence lengths. Again, arrows are included to show whether each disparity grows or shrinks in the simulated evidence-based model. This table shows very different results for the effects of evidence-based sentencing than were observed in the decision to incarcerate

[46] These results do not change substantively when coefficients from the .5 or .6 threshold models are used in place of the .4 threshold model.

comparisons. Here, every statistically significant differences shows *less* disparity in the evidence-based scenario; the disadvantages for black, Hispanic, male, less educated, unemployed, and financially unstable offenders (based on reliance on government assistance) are all statistically significantly smaller in simulated evidence-based sentence lengths. Furthermore, some of these differences are quite substantial. The average marginal effects for black and Hispanic offenders shrink by 68% and 72%, respectively, and the average marginal effect for males decreases by 54%. The significant reductions in average marginal effects for SES indicators are somewhat smaller but still meaningful, ranging from 7% to 44%.[47] The effects for Asians, offenders who either completed 10th grade or had some post-graduate education, and offenders with lower financial problems scores do not differ substantially between the actual and simulated evidence-based sentence length models.

Circling back to the final research question in the study, these results provide preliminary evidence that evidence-based sentencing does have the potential to increase some sociodemographic disparities in punishment, primarily through differences in the likelihood of incarceration. Disadvantages for black, Hispanic, less educated, and financially unstable offenders are larger in the simulated evidence-based decision to incarcerate than in the decision to incarcerate in Connecticut's current system. This suggests that structural sources of disparity, such as LSI-R scores used in an evidence-based sentencing scheme, have the potential to generate as big or even bigger disparities

[47] The magnitude of differences tends to be smaller, but only few substantive conclusions change when coefficients from the .5 or .6 threshold model are used instead of the .4 threshold model. Notably, the effect of reliance on social assistance is larger in simulated evidence-based sentence lengths than in actual sentence lengths when the .5 threshold model is used, and the effects of both employment and reliance on social assistance are larger in evidence-based sentence lengths than in actual sentence lengths when the .6 threshold model is used.

in punishment than judicial discretion alone. In stark contrast, disparities are almost uniformly lessened in simulated evidence-based sentence lengths. Though disparities in simulated sentence lengths are observable, they constitute a much smaller amount of inequality than disparities in Connecticut's current sentence lengths.[48] All in all, these findings suggest that an evidence-based sentencing scheme has the potential to exacerbate social inequality in Connecticut by creating larger differences in incarceration rates between groups, but once incarceration has been assigned, the addition of risk assessment scores into the sentencing process has less potential to worsen existing disparities and may even reduce them to a degree.

[48] It is worth noting again that this is true even when alternative evidence-based specifications that result in higher evidence-based disparities are considered.

CHAPTER 9: DISCUSSION

Research on social inequality in criminal punishment has maintained a position at the forefront of courtroom literature for decades, spurred on by countless studies that identify sociodemographic disparities in sentencing (Baumer, 2013; Daly & Bordt, 1995; Mitchell, 2005; Ulmer, 2012; Wu & Spohn, 2009; Zatz, 2000). The advent of innovations aimed at reforming sentencing processes in the criminal justice system introduced new sources of disparity in punishment, opening the door for scholars to examine the effects of sentencing reform on observed disparities (Bushway & Piehl, 2007; Ulmer et al., 2007; Zatz, 2000), but newer innovations such as evidence-based sentencing have not been evaluated in this manner.

Evidence-based sentencing is lauded by supporters as a data-driven approach to risk evaluation and sentencing, one that has the potential to help judges more effectively differentiate between high-risk and low-risk offenders, identify offenders who would benefit most from rehabilitative programming, and reduce reliance on incarceration (Hyatt et al., 2011; Monahan & Skeem, 2015; Warren, 2007). At the same time, critics assail the practice for its use of profiling and possibly discriminatory consideration of factors that historically correlate with members in marginalized groups (Hannah-Moffat, 2005, 2013; Starr, 2014b). Risk assessments, while perhaps more precise than clinical judgments, may still generate substantial margins of error in risk prediction (Berk & Bleich, 2014; Harcourt, 2007; Hart et al., 2007). Lastly, and most relevant for this study, scholars argue that evidence-based sentencing will increase social inequalities in punishment (Hannah-Moffat, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b,c).

This study seeks to inform the efficacy of this final criticism, capitalizing on recent interest from the state of Connecticut in understanding the potential effects of introducing the LSI-R into its sentencing system. Connecticut, which does not currently have sentencing guidelines, is well-suited for this study of evidence-based sentencing. Broad offense categories allow ample room for judicial discretion that can be shaped by risk evaluation, and the state's familiarity with (and openness to) risk assessment tools helps maximize the study's potential utility.

At its core, the study evaluates disparities based on race, ethnicity, gender, and SES in three settings: Connecticut's current sentencing scheme, the LSI-R, and a simulated evidence-based sentencing scenario. This chapter discusses results from these three setting evaluations, as well as how the results can be combined conceptually to inform how the introduction of the LSI-R into sentencing decisions may be expected to alter patterns of sociodemographic disparity in punishment. It concludes with a consideration of study limitations, directions for future research, and policy implications.

Discussion of Findings

In order to properly evaluate whether evidence-based sentencing has the potential to change sentencing patterns in Connecticut, the first aim of the study focused on characterizing patterns of disparity according to race, ethnicity, gender, and SES (as captured by educational attainment, employment status, and financial situation) in Connecticut's existing sentencing structure. Using sentences issued to offenders during a three-year period, analyses were conducted to first identify aggregate sociodemographic disparities in the decision to incarcerate and sentence length, then to consider how much of those aggregate disparities could be explained by legal and case processing factors and later factors that tap into offenders' risk of recidivism as well.

Results from negative binomial hurdle models showed some interesting trends. Patterns of disparity differed substantially between the decision to incarcerate and sentence length. In the decision to incarcerate, aggregate disparities by gender and SES generally disadvantage male and low-SES offenders, though financial instability appeared to serve as an advantage. Accounting for legal, case processing, and risk factors reduced the gender, employment, and financial problems gaps but exposed a disadvantage for the most highly educated offenders. Meanwhile, the aggregate analysis unexpectedly showed little disparity according to race and ethnicity, and the addition of legal/ case processing characteristics and risk factors actually revealed a significant advantage for black and Hispanic offenders in the decision to incarcerate.

Among sentence lengths, however, most of the social inequality patterns take another form. Black, and Hispanic offenders were disadvantaged in the aggregate and maintain a disadvantage even after legal, case processing, and risk factors are controlled. This trend is more consistent with traditional theory, which posits that racial and ethnic minorities will be perceived as more blameworthy, more risky and dangerous, and more able to endure the harsh conditions of incarceration (Albonetti, 1991; Steffensmeier et al., 1998). For male, low-educated, and unemployed offenders, aggregate disadvantages were explained partially by legal/case processing characteristics and factors that estimate risk of recidivism. Disparities by financial situation also depended upon the indicator used; reliance on government assistance is advantageous in the aggregate, and a difference persists after accounting for legal/case processing and risk factors, but no relationship between the financial problems scale and sentence length was observed.

The first conclusion to be drawn from these findings is that sentencing disparities exist in Connecticut in the absence of an evidence-based sentencing scheme. In the aggregate, disparities in every sociodemographic characteristic considered were evident in one or both of the decision to incarcerate and sentence length. Even after accounting for legal and case processing factors, several disparities persisted, and in some cases the residual disparities were still large enough to be substantively meaningful. These results are consistent with expectations; as noted throughout this study, prior research supports the notion that sociodemographic disparities in punishment exist at least some of the time, and in some contexts, even after accounting for legally relevant offender and offense characteristics (Baumer, 2013; Daly & Bordt, 1995; Mitchell, 2005; Ulmer, 2012; Wu & Spohn, 2009; Zatz, 2000).

As described by theorizing from the behavioral economics and sociolegal literatures, these persistent disparities may be the result of judges using heuristics to make sentencing decisions under conditions of limited time and resources (Scholz, 1983). In high-pressure situations, judges turn to stereotypes and other cognitive shortcuts that make decision-making more efficient but also open it up to errors that may systematically disadvantage certain groups of individuals (Tversky & Kahneman, 1974). When judges use heuristics to inform their perceptions of offenders' blameworthiness and risk, the process can result in biased perceptions and ultimately generate sociodemographic disparities in punishment (Albonetti, 1991; Steffensmeier et al., 1998). Even after accounting for legal, case processing and risk factors, these disadvantages will persist.

However, the finding in this study that black and Hispanic offenders are less likely to receive incarceration sentences than whites after controlling for legal, case processing factors, and risk does not sync well with popular sentencing theorizing, which typically relies on cultural stereotypes of racial and ethnic minority individuals as violent, unstable criminals to explain why judges are more likely to issue those groups harsher punishment at sentencing (e.g. see Steffensmeier et al., 1998). One potential explanation for this observation is provided by Bernstein and colleagues (1977), who discuss the potential for a cultural acclimatization effect following their discovery that white defendants were sentenced more severely than black or Hispanic defendants in a sample of male felony offenders in New York (see p. 753). Because deviant behavior is often viewed as acceptable in non-white subcultures, they argue, familiarity with non-white subcultures normalizes non-white crime in the eyes of judges and prosecutors, making it appear less offensive. White subcultures are less tolerant of deviant behavior, on the other hand, so white crime is viewed as a violation of both the law and white cultural norms, and it consequently earns more severe punishment for white offenders. Though this phenomenon cannot explain specifically why racial and ethnic effects would differ between the decision to incarcerate and sentence lengths, it does offer a meaningful counterpoint to the expectations of attribution and focal concerns theories and provides some context for why minority offenders may be treated relatively leniently in some cases. Judges may view crimes by white offenders as more deviant and therefore more deserving of harsher punishment.

This finding may additionally reflect unique racial and socioeconomic dynamics in the state of Connecticut. Connecticut's population is nearly 80% white, and though the state is relatively wealthy as a whole, income inequality between the top and bottom earners is substantial (Sommeiller, Price, & Wazeter, 2016; U.S. Census Bureau, 2017). With such little racial and ethnic diversity, it stands to reason that a nontrivial part of Connecticut's poorest proletariat is comprised of white residents. The existence of a white underclass, whose pathology includes the same limited social and economic resources that have historically characterized minority populations, would help explain why white offenders receive harsher punishments, on average, in the criminal justice system. Subject to the same marginalization as minority offenders in other more diverse areas of the U.S., many white offenders may be punished for their deviance with sentences that exceed those of other racial and ethnic groups.

In the context of this study's larger goals, these results altogether highlight the need to identify the correct baseline levels of disparity before assessing how evidence-based sentencing might change them. The disparities observed in Connecticut's existing sentencing scheme serve as reference points with which to compare disparities in evidence-based sentencing. Critics use sociodemographic disparities in actuarial risk assessment instruments to suggest that evidence-based sentencing would worsen social inequality, but this evidence only demonstrates that there are likely to be disparities in evidence-based sentencing, not that there are likely to be *larger* disparities in evidence-based sentencing. In fact, the observation that disparities are already noticeable in Connecticut shows that introducing actuarial risk assessment in the state even has some ipso facto opportunity to reduce social inequality in punishment.

Though racial and ethnic disparities are more often the focus of sentencing scholarship (see Baumer, 2013; Mitchell, 2005; Pratt, 1998; Spohn, 2000; Ulmer, 2012;

Zatz, 2000), the disadvantages for low-SES offenders were among the largest disparities observed in this study's analysis of Connecticut punishment outcomes. Evidence from prior literature generally supports the notion of a prominent role for SES in sentencing (Chiricos & Bales, 1991; Franklin, 2017; Miethe & Moore, 1985; Wooldredge, 2010), but research on the topic is limited due to difficulties measuring SES in official records. SES is typically captured using only one of its multiple dimensions if it is captured at all. Results from this study, which boasts a fairly robust set of indicators representing all three of the widely accepted components of SES, suggest that the way SES is captured can affect substantive conclusions made about its relationship with legal variables, case processing characteristics, and sentencing. Educational attainment, employment status, reliance on government assistance, and financial problems all yielded somewhat different trends across model specifications, underscoring the importance of measuring SES clearly and completely in order to better understand the role SES plays in determining punishment.

Switching gears, the unique risk factors included in this sentencing evaluation accounted for relatively little portions of remaining disparities. Given that sociolegal theorizing frequently emphasizes the role of risk perceptions in both sentencing decisions themselves and in disparities in those sentencing decisions (Albonetti, 1991; Spohn, 2009; Steffensmeier et al., 1998; von Hirsch, 1976), it is perhaps surprising that individual elements of risk captured in the LSI-R did not explain more of the residual sociodemographic disparities in sentencing. It may be the case that the range of legal and case processing factors included in this study sufficiently captured variation in punishment that could otherwise be attributed to differential risk perceptions. A second possibility is that racial, ethnic, gender, and SES group memberships do uniquely affect judicial perceptions of offender risk, but they do so in ways that are not systematically related to the latent risk characteristic measured by the LSI-R.

Though risk factors did not explain large portions of sociodemographic disparity, some of them did have their own independent effects on punishment in Connecticut. Age had only a slight impact on the two sentencing outcomes, and marital status likewise had little effect. Having a previous conviction in the dataset, however, had a substantial effect on the odds of incarceration (not sentence length) even after accounting for number of prior convictions. Though the measure is included in analyses to prevent offenders who appear multiple times in the dataset to have an outsized effect on results, the measure may also serve as a rough indicator of offending frequency. If offenders who cycle through the criminal justice system repeatedly in relatively short amounts of time are treated more harshly than offenders who manage to last years before being convicted again, this variable may have captured some of that effect.

Scores on criminogenic domains included in the LSI-R also have independent effects on punishment. While higher scores on most domains are associated with higher likelihoods of incarceration and/or longer sentence lengths, the direction of the Emotional/Personal domain effect in particular appears to depend on the punishment outcome. Higher Emotional/Personal scores are associated with longer sentences but lower odds of incarceration. As mentioned in the presentation of results (see also Appendix B), the items in the Emotional/Personal domain focus on offenders' mental health, tapping into diagnoses, treatment, and severity of mental illness. Though mental illness may be an indicator of higher recidivism risk, judges may also view it as a reason to issue a non-incarceration sentence, either because the offender is perceived as less responsible for his/her crimes or because incarceration would interfere with any ongoing treatment. These two separate mechanisms may explain why the Emotional/Personal domain score is associated with both more severity in sentence lengths but more leniency in the decision to incarcerate.

The domains for which higher scores are most notably associated with more severe sentences in Connecticut even without the use of evidence-based sentencing are the Accommodation, Companions, and Attitudes/Orientations domains, though the magnitude of most associations are modest. Accommodation scores are most impactful for the decision to incarcerate, while the effects of the Companions and Attitudes/Orientations scores are centered in sentence length. These are all rather interesting relationships to ponder.

The Accommodation domain estimates the quality and stability of the offender's living situation, as well as whether the offender lives in a high-crime neighborhood. These characteristics likely correlate with socioeconomic status (a notion supported by findings from Phase 2 of this study), but even after controlling for educational attainment, employment status, and financial situation, the Accommodation domain continued to have a significant association with incarceration. This suggests that there is some other aspect of offenders' residential situation that influences judges' sentencing decisions. Perhaps residential instability and residence in a high-crime neighborhood behave as proxies for gang affiliations (Sampson & Wilson, 1995), which may lead prosecutors to offer less favorable plea deals and judges to issue harsher sentences. Perhaps frequent residential changes and an unsatisfactory current living situation indicate homelessness, to which judges may respond by using jails as either temporary shelters or as warehouses for managing transient "rabble" (Irwin, 1985).

The Companions domain is the first of two domains that most strongly impacted sentence lengths for incarcerated offenders (though again, the incremental effect is relatively modest). This category measures offenders' social isolation, criminal acquaintances/friends, and non-criminal acquaintances/friends. The link between criminal peers and deviance is well-established (Gendreau et al., 1996; Elliott et al., 1985; Thornberry et al., 1994), but unless judges routinely ask adult offenders about their relationships with peers, this does not provide sufficient explanation for why offenders who have mostly criminal friends or few friends at all would receive harsher punishment at sentencing. Instead, it may be that this domain taps into gang affiliations, which likely lead to more criminal friends/acquaintances and fewer non-criminal friends/acquaintances. Alternatively, it may indicate a lack of involvement in prosocial activities (recreational sports, hobbies, etc.), which often introduce people to non-criminal peers.

The other most influential domain predictor of sentence length is the Attitudes/Orientations domain, which captures offenders' sentiments about crime, convention, and punishment in the criminal justice system. If these sentiments are expressed by offenders either verbally or nonverbally during courtroom appearances, it is easy to envision how they could come to be associated with longer sentences. Prior research has demonstrated the importance of nonverbal courtroom cues such as demeanor (Frazier, 1979; Hedderman, 1990) and showing remorse (Harrel, 1981; Robinson et al., 1994), so perceived disrespect toward the justice system and its actors may come off negatively to a prosecutor or a judge, as may perceived commitment to a deviant lifestyle. Antisocial attitudes are also correlated with recidivism (Gendreau et al., 1996), so prosecutors and judges may also view offenders who express more negative attitudes toward the conventions of the criminal justice system as riskier.

The second general aim of the study was to characterize disparities in the LSI-R and in evidence-based sentencing based on LSI-R scores. Critics of evidence-based sentencing contend that the practice will deepen unwarranted punishment disparities because items on actuarial risk assessments correlate with sociodemographic characteristics, allowing indirect pathways for biases to creep into sentencing decisions. Phase 2 of this study sought to test the foundations of that argument by first assessing disparities in LSI-R composite and domain scores.

Disparities in LSI-R composite scores according to race and ethnicity are generally modest. Asian offenders receive substantially lower scores than any other racial group (and the disparity is spread across the majority of criminogenic domains), but the margins for black and Hispanic offenders compared to whites are both close to one point. In these cases, though, composite scores conceal a variety of meaningful relationships between risk elements and race/ethnicity. Black and Hispanic offenders receive higher average scores in eight of the ten total domains, with the largest gaps centered in Criminal History and Education/Employment, but the two remaining domains counteract nearly all of the disadvantage: white offenders score much higher in the Alcohol/Drug and Emotional/Personal categories than blacks or Hispanics. The pattern for the Alcohol/Drug domain stands in direct contrast to cultural stereotypes of black and

Hispanic individuals as drug users (Curry & Corral-Comacho, 2008; Gibbs, 1988; Richey-Mann et al., 2006).[49]

Though consistent with prior research (e.g. Chenane et al., 2015), it is a little less clear how the patterns observed in the Emotional/Personal category fit with current scholarly understanding of race/ethnicity and mental health. There is little evidence that mental health problems are more prevalent among one racial/ethnic group or another in correctional settings (James & Glaze, 2006), but given that several items in the Emotional/Personal domain focus on receiving mental health treatment rather than just having mental health problems, it is possible that much of the racial/ethnic gap in Emotional/Personal scores is due to differences in either access to mental health care or culture-specific attitudes toward mental health treatment (U.S. Department of Health and Human Services, 1999).

The LSI-R composite score difference between male and female offenders likewise averages out remarkably close to zero, which is at least not discordant with previous research's mixed conclusions about gender differences in risk scores (Andrews & Bonta, 2003; Lowenkamp et al., 2001; Mihailides et al., 2005). However, the composite similarity again masks more variable patterns of disparity in individual LSI-R domains. As expected, men score higher on the Criminal History domain. They also score higher in Leisure/Recreation, Companions, Alcohol/Drug, and Attitudes/Orientations. These results match prior work on disparity and risk assessments (Manchak et al., 2009; Mihailides et al., 2005). Trends in many of the domains in which women appear "riskier"

[49] The researcher initially posited that this relationship may have been tied to crime type: white offenders are more likely to have convicted of a drug crime and are therefore more likely to have alcohol or drug problems. Supplemental analyses indicate that though white offenders more often received drug convictions (33% vs 27% for blacks and 28% for Hispanics) in this sample, the relationship between race/ethnicity and Alcohol/Drug domain scores is not accounted for by the commission of drug crimes.

mirror findings from other criminological and feminist scholarship as well. Higher female scores on the Emotional/Personal domain are consistent with research that finds higher rates of mental health concerns among women than men in incarcerated populations (James & Glaze, 2006). Higher female scores on the Family/Marital domain are consistent with the tendency for females to be led into criminal activity by male significant others or family members (e.g. Miller, 1986), while higher female scores on the Financial and Accommodations domains support the narrative that female offenders are often worse off economically than their male counterparts (Chesney-Lind, 1997).

Among LSI-R composite scores, the largest disparities emerge for SES groups. Even after accounting for the fact that educational attainment, employment status, reliance on government assistance, and financial problems are directly captured in the LSI-R, substantial gaps based on SES indicators are apparent. The gap between the highest- and lowest-education offenders reaches upwards of 12 points, the gap between full-time employed and unemployed offenders is almost 11 points, reliance upon government assistances increases average risk scores by 6.3 points, and offenders with very unsatisfactory finances have scores that are 13.9 points higher than offenders with very satisfactory finances. Subsequent domain-specific analyses indicate that these SES disparities are spread across every single LSI-R domain; disadvantages for uneducated, unemployed offenders who rely on government assistance and have very unsatisfactory finances appear across the board.

Altogether, these are rather intriguing findings in light of the fact that many scholars choose to focus on racial and ethnic rather than socioeconomic disparities in risk evaluation. Some scholars advocate specifically for wide-ranging tools like the LSI-R, concluding that risk scores produced using a wider set of criminogenic factors predict recidivism better and tend to be less correlated with race and ethnicity (Berk, 2009; Skeem & Lowenkamp, 2015). The results from this study, on the other hand, suggest that racial and ethnic differences in risk evaluation would not necessarily be the largest and most egregious source of social inequality in evidence-based sentencing. Risk score disadvantages based on SES may instead hold that title. The LSI-R, this study suggests, is not kind to low-SES offenders.

The next section of the study evaluated whether or not these apparent disparities in risk evaluation may translate into disparities in evidence-based sentencing outcomes. Analyses first established that there is indeed a meaningful relationship between LSI-R composite scores and punishment in Connecticut's current system even after controlling for legal and case processing factors, which lent some credence to the study's crucial assumption that judges conduct informal risk evaluation even in the absence of evidence-based sentencing and that their risk evaluation taps into the same underlying risk trait that the LSI-R measures. If LSI-R scores did not have an observable relationship with punishment outcomes, this assumption would not able to hold water. The second step in the process created a simplified simulation of evidence-based sentencing by assigning incarceration and sentence lengths to all cases using legal/case processing factors and LSI-R scores as determinants. Different probability thresholds were used to determine which cases were assigned evidence-based incarceration, and as the results showed, these different thresholds had a substantial effect on both the portion of cases that were assigned incarceration and sentence lengths. After considering the distributions created using each threshold rule, the researcher concluded that the simulation produced under the lowest probability threshold (.4) had the most similar distribution to actual sentencing outcomes and was therefore most appropriate for assessing disparities in the simulated evidence-based sentencing scenario.

Sociodemographic disparities in the resultant simulation were noticeable. In the simulated decision to incarcerate, black, male, and low-SES offenders were at a large disadvantage. In simulated sentence lengths, race/ethnicity had less of an effect, but male and low-SES offenders were assigned longer sentences. Given that the gender difference in LSI-R composite scores was almost non-existent and the racial/ethnic differences were relatively small as well, the disparities for black and male offenders observed in punishment outcomes are likely due far more to differences in legal and case processing characteristics than to differences in LSI-R scores. Much of the differences for low-SES offenders, however, can be linked back to disparities in the LSI-R, where low-SES offenders racked up considerably higher risk scores regardless of which SES indicator was used to assess.

The substantive conclusions were not particularly sensitive to the threshold decision rule used to assign evidence-based incarceration. Though the .4 threshold model served as the primary model, the .5 and .6 threshold models also led to sizeable disadvantages for black, male, and low-SES offenders. What differed between the three conditions was the magnitude of the disparities observed. As the threshold increased, the magnitude of disparities in the evidence-based decision to incarcerate model grew slightly, and the magnitude of disparities in the evidence-based sentence lengths model grew dramatically. This trend indicates that disparities in the simulated evidence-based scenario created in this study are concentrated at the upper end of the evidence-based incarceration probability distribution, where legal/case processing factors and risk scores tend to signal for more punitive punishment.

Ultimately, this study sought to address one of the more contentious criticisms of evidence-based sentencing, that it will result in increased sociodemographic disparities in punishment. Therefore, the final research aim of the study was to assess potential changes in disparity between Connecticut's current sentencing scheme and an evidence-based sentencing approach. To do this, the study juxtaposed comparable coefficients from the aggregate disparity model of Connecticut's actual sentencing system with the aggregate disparity model of the simulated evidence-based scenario.

The findings from this comparison do not conform to critics' predictions quite as consistently as they could. However, in the decision to incarcerate, there did appear to be several disparities that were aggravated in the simulated evidence-based sentencing model. In support of scholars who express concern that racial and ethnic disparities in punishment will grow if risk assessments are incorporated into sentencing (Hannah-Moffat, 2005, 2009, 2013; Monahan & Skeem, 2015; Starr, 2014b,c), this study did find that both black and Hispanic offenders have a larger aggregate disadvantage in simulated evidence-based sentences than in Connecticut's current sentencing scheme. The difference in magnitude for the black effect in particular was sizeable, with an odds ratio in the simulated model that was over three times larger. For males, the gap was smaller in simulated evidence-based sentences, which can again be interpreted as unsurprising given that the average gender difference in LSI-R composite scores was negligible. Among SES categories, it was educational attainment and both indicators of financial situation that emerge as aspects of SES for which disparities may be increased in an evidence-based sentencing scheme; employment status appeared to have a slightly smaller (though still large) role in determining punishment in the simulated model. Among sentence lengths, on the other hand, this study provides little evidence that an evidence-based sentencing scheme should be expected to exacerbate social inequality. Where significant differences between actual and simulated disparities existed, the disparities were all smaller in the evidence-based scenario. Though the magnitude of differences varied according to the incarceration probability used in the simulated model, substantive conclusions were not sensitive to this change.

These results are a tentative indication that concerns about an increase in social inequality following the implementation of an evidence-based sentencing scheme may be well-founded. Though larger disparities weren't apparent across the board, socioeconomic disparities in the LSI-R translated into differences in simulated incarceration rates that generally disadvantaged black, Hispanic, and low-SES offenders more than judicial discretion did in Connecticut's actual sentencing scheme.

Placing this study's findings in the context of overarching themes in criminal justice research and policy leads to several worthy points of discussion. First, the consistency of disadvantages for low-SES offenders observed in this study invites commentary about how low-SES offenders are viewed and treated in the justice system. SES disparities are readily observable in Connecticut's current scheme, a fact that corresponds well with previous research on the relationship between SES and sentencing (D'Alessio & Stolzenberg, 1993; Chiricos & Waldo, 1975; Kramer & Ulmer, 2009; Miethe & Moore, 1985). Of course, some differences in punishment are likely due to behavioral differences between high- and low-SES offenders, as evidenced in this study by the small amount of SES disparity accounted for by legal and case processing factors, but residual disparities indicate that there is more to the story. Conflict theories of criminal justice posit that this relationship is the result of the social elite using the justice system to subdue marginalized group in society (Chambliss & Seidman, 1971; Quinney, 1973), but this study's evaluation of disparities in LSI-R composite and domain scores offers an additional explanation for why this pattern may emerge.

The consistency of higher scores on the LSI-R for low-SES offenders in this study is striking, and it raises concerns about just how closely the criminal justice system's concept of risk is intertwined with socioeconomic status. Hannah-Moffat (1999: 71) describes risk as "ambiguous, fractured, and flexible", a characterization which hints at ample room for subjectivity and bias in the risk evaluation process. Conceptualizations of risk tend to rely on both cultural and moral evaluations that could be tied, even unintentionally, to sociodemographic characteristics such as SES. Even though the LSI-R is praised for its inclusion of a variety of seemingly neutral criminogenic domains, low-SES offenders received higher scores in every domain captured in the LSI-R in this study, essentially turning actuarial composite risk scores into a proxy for SES. Under an evidence-based sentencing scheme, where actuarial assessments are intended to shape judicial discretion by informing perceptions of offenders' risk, it is easy to see how the LSI-R's association with SES could translate directly into punishment disparity. Even without an evidence-based sentencing system in place, though, perceptions of offender risk still play an important role in punishment determinations. If the justice system's perceptions of offender risk are otherwise built on factors that correlate highly with SES, then differential assessments of risk may serve as another mechanism through which SES disparities emerge in punishment contexts in the absence of actuarial risk assessment. In this evaluation, which pits punishment disparities borne from judicial discretion against disparities from structural forces like risk assessments, low-SES offenders come out on the bottom no matter how you slice it.

Close scrutiny of the domains in the LSI-R also brings to light an interesting pattern whereby factors that determine an offender's risk score are in other sentencing contexts regarded as unwanted, unwarranted influences on punishment. The popularity of risk assessment tools like the LSI-R, coupled with a lack of intensive discussion about the individual items used to construct them, suggests that by virtue of their status as significant predictors of recidivism risk, actuarial risk assessments can be treated as legal factors that can effectively be used to justify disparities in sentencing. Resistance to challenging actuarial risk assessments as a valid sentencing factor may spring from the comfortability of placing confidence in numbers and scores. Policymakers, and people more generally, enjoy using numbers to justify their decisions; numbers tend to carry with them an air of factualness and an implicit authority that wards off criticism (Porter, 1996). Hannah-Moffat (2013: 277) wrestles with this idea in the context of actuarial risk assessment, noting that "Risk scores impart a sense of moral certainty and legitimacy into the classifications they produce." Undoubtedly, the fact that risk scores are numbers 'backed by science' makes them both more palatable and more readily accepted. As a result, the objectivity and neutrality of the items used to construct them are more often assumed than tested.

But these same characteristics that are deemed suitable in the context of predicting risk of recidivism evince stronger negative reactions when used as direct explanations for punishment. In this study's evaluation of disparities in Connecticut's current sentencing scheme, several significant and meaningful disparities were observed. But while the focus was on disparities according to sociodemographic characteristics, specifically race, ethnicity, gender, and SES, disparities according to domain scores from the LSI-R emerged as well. The Accommodations, Companions, and Attitudes/Orientations domains in particular had significant and nontrivial associations with punishment. In a traditional sentencing study, these relationships would be characterized as "unwarranted disparities"; scholars routinely lament the existence of punishment differences according to characteristics such as educational attainment or neighborhood of residence (Franklin, 2017; Wooldredge, 2007). The implication from the tone of such studies is that these disparities are neither desirable nor justifiable.

Thus, the perceived acceptability of linking social characteristics to sentencing may depend on whether or not the relationships are observed through the lens of risk evaluation. If the association is framed as a direct relationship, without risk assessment serving as an intermediary, it is difficult to view the association as anything but unwelcome. On the other hand, if the association is framed as indirect, such that factors such as having criminal relatives or little organized activity are connected to sentencing decisions only through their relationship with risk, it becomes more tolerable. It may be that this indirect pathway allows decision-makers to mentally compartmentalize social and other extralegal factors as predictors of *risk* rather than predictors of *punishment*, a subtle distinction that nevertheless could allow decision-makers to cognitively divorce themselves from the moral complication associated with using extralegal characteristics to make evidence-based punishment decisions.

To the researcher's knowledge, there has also been little political or scholarly discussion about the implications of translating an actuarial instrument designed for correctional settings to the courtroom in its entirety. In *Malenchik v. State of Indiana* (2010), a defendant challenged the use of his high scores on two actuarial assessments (one of which was the LSI-R) as aggravating circumstances at sentencing, arguing that the practice was discriminatory on a variety of grounds. Malenchik's arguments were summarily rejected by the Indiana Supreme Court, and in its decision, the Court concluded that "an evidence-based tool such as the LSI-R may be utilized in the sentencing process if employed consistently with its proper purpose and limitations." (*Malenchik v. State of Indiana, 2010:* 3).

What are the proper purpose and limitations of the LSI-R? The tool was originally designed to assist practitioners in the correctional system with assigning supervision levels and lengths to offenders (Andrews, 1982). It is relatively easy to see how a tool used for this function only can be applied to judicial decision-making: just as corrections staff may impose tighter, longer supervision on riskier offenders, judges may impose incarceration and longer sentences on riskier defendants. However, consistent with the principles of the RNR model on which the tool is based, amendments to the instrument shifted part of its focus to identifying service and treatment provisions. It is now intended to measure dynamic risk and serve as an aid for risk reduction rather than simply risk management. Assessment results are used to target high-risk offenders in particular for rehabilitative programs that can help reduce their risk of recidivism, and in some cases, the commitment to rehabilitation may result in what is perceived as more "lenient" treatment.

In the sentencing system, it is unclear whether there is a viable role for risk assessments to help judges contribute to risk reduction rather than risk management goals only. It is difficult to reconcile using scores on the LSI-R to justify harsher sentences (without correspondingly more service provisions) with the RNR model's emphasis on rehabilitation rather than deterrence. The tendency to treat risk scores as a barometer of sentence severity rather than as a signal for rehabilitative intervention options creates ample opportunity for evidence-based sentencing to become a tool for justifying selective incarceration rather than for preventing recidivism (Garland, 2003; Hannah-Moffat, 1999). What is therefore needed is more in-depth consideration of whether/how the LSI-R and other similar correctional risk assessment instruments can be adapted to guide sentencing decisions without compromising the integrity of the service-oriented principles on which they are based.

The final theme that the results in this study speak to is the tension between effectiveness/efficiency and fairness in evidence-based sentencing. The LSI-R is a well-validated risk prediction tool that performs relatively well in 'competition' with clinical risk judgments, which makes it an asset to the criminal justice system as a means of improving precision in judicial risk evaluations conducted at the time of sentencing. However, its predictive validity comes with costs, and one of those costs is a nearly systematic disadvantage for offenders who are often marginalized in other non-criminal justice settings. This study's findings identified disparities in a simulated evidence-based sentencing scenario that disadvantaged black, Hispanic, male, and low-SES offenders, and in several of those cases, the disadvantages exceeded those observed in Connecticut's actual sentencing scheme. Thus, in the pursuit of effective risk prediction and more efficient use of resource-intensive punishments, the criminal justice system compromises some of the fairness and equity that it promises to espouse.

It is unrealistic to assume that either of these objectives will ever be permitted to entirely dominate the other; a maximally precise sentencing scheme with complete disregard for equity would dramatically undermine the legitimacy of the entire criminal justice system, while a perfectly equitable sentencing system without any precision in risk prediction would be of little value to it. Therefore a balance between the two ideals must be struck, and the task is to identify the appropriate balance between predictive validity and equity. How much inequality in punishment is society willing to tolerate in order to advance public safety and use limited criminal justice resources more efficiently? This is a difficult ethical question, one for which the answer will likely depend on who is giving it. Even so, it is one that scholars, policymakers, and practitioners must all engage with if evidence-based sentencing is to become a mainstay in criminal justice policy.

Limitations

This study, of course, has its limitations. First, the most unique aspect of this project, the simulation of evidence-based sentences, brings with it a unique set of drawbacks. Because the evidence-based sentence building procedure uses the existing relationship between risk and sentences in Connecticut's current scheme to estimate the relationship between risk and punishment in a simulated evidence-based scenario, the study assumes two things: first, that LSI-R risk evaluations and judges' informal risk evaluations tap into the same underlying risk characteristic for each offender, and second, that the aggregate impact of risk on punishment would not change if risk assessments were incorporated into sentencing outcomes. The second assumption in particular is tenuous, as some scholars suggest that evidence-based sentencing is a method for turning the spotlight onto public safety and increasing the relative impact of recidivism risk on sentencing (Etienne, 2009; Marcus, 2009; Warren, 2007). If this is the case, and incorporating risk assessments into sentencing leads to a greater role for risk in sentencing decisions, then the current study's estimates may be a conservative estimate of potential differences in disparities between Connecticut's current system and an evidence-based system.

This conservative estimation would have differing effects on the substantive conclusions from this study for the decision to incarcerate and sentence lengths. For the decision to incarcerate, this study provided evidence that disparities may increase under an evidence-based sentencing scheme. If the disparity estimates in the simulated evidence-based scenario are low, then disparities in an actual evidence-based sentencing system would be expected to be even larger, amplifying the gap between evidence-based disparities and disparities in Connecticut's current scheme. This effect would only bolster the general conclusion that evidence-based sentencing has the potential to increase disparities in the decision to incarcerate even more. Conversely, for sentence lengths, results in this study indicated that for most sociodemographic factors, evidence-based sentences generated smaller disparities in sentence lengths than Connecticut's current scheme does. In this case, if estimates of disparity in the simulated evidence-based scenario are low, the study is overestimating the gap in disparity between the two sentencing systems, and evidence-based disparities would be expected to be more similar to (or possibly even larger than) disparities in Connecticut's existing system than results suggested. The general conclusion that evidence-based sentencing has the potential to actually decrease disparities in incarceration lengths would be undermined.

Relatedly, the evidence-based sentencing scenario is modeled for a sample of sentenced offenders. All of the offenders have been charged, some of them were likely detained pretrial, many of them have negotiated plea deals, and those who did not agree to a plea deal were convicted at trial. However, current perspectives on criminal case processing (Kutateladze et al., 2014; Wooldredge et al., 2015) emphasize the importance of these earlier case processing outcomes for punishment determinations; small disparities at several different case processing decision nodes can accumulate as each defendant moves through the system, ultimately effecting large differences in punishment. Moreover, if implemented, evidence-based sentencing may have consequences on some of these earlier case processing decisions. It is possible that introducing risk assessments into sentencing outcomes would alter the decision-making patterns of court actors other than judges in ways that would have downstream effects on a) which offenders are convicted and sentenced and b) disparities in sentencing outcomes. For example, prosecutors have vast discretionary authority in processing cases (Johnson, King, & Spohn, 2016), and risk assessment scores may become a piece of information that prosecutors use as leverage to negotiate more favorable plea deals. These possibilities, and their influence on sentencing patterns, are not modeled or accounted for in the current study.

There are also two concerns related to sample selection in this study that must be considered. First, the use of a sentenced sample introduces the potential for selection bias due to the omission of cases that did not result in conviction. It is likely that the group of cases selected into conviction and sentencing differs from the full population of cases that enter into the criminal justice system. Estimates derived from modeling punishment outcomes among convicted offenders may therefore be biased (Berk & Ray, 1982; Hagan & Parker, 1985). Previous research has demonstrated that bias can indeed be introduced into sentencing analyses by excluding cases that do not result in conviction and equally importantly, that the effects of other predictors on punishment can differ when bias is present (Zatz & Hagan, 1985). Though many scholars acknowledge this selection concern, it is less often addressed well empirically (see Bushway et al., 2007; Johnson, 2014 for further discussion). The current study does not account for this possibility either. In the future, it would be ideal to model selection into the convicted sample, in order to estimate how this selection process may influence obtained disparity estimates.

The potential for selection bias is additionally ushered in by the number of cases that were excluded from analysis due to a lack of LSI-R information. Descriptive statistics showed a marked difference in several sociodemographic and legal characteristics between the cases that were and were not matched with a complete LSI-R risk assessment. These differences indicated that more serious cases, as indicated by more/longer incarceration sentences and longer criminal histories, were less likely to be matched. Because more serious cases were less likely to be matched to a complete risk assessment, estimates of disparities in both Connecticut’s actual sentencing system and in the simulated evidence-based sentencing scenario may have been affected. A consequence of this is that the external validity in the study is limited to sentenced felony or misdemeanor cases in which a complete LSI-R risk assessment is administered under Connecticut’s current LSI-R administration schedule. In other words, if Connecticut changed its risk assessment policy and began administering the LSI-R assessment to all offenders in the state, the analyses conducted in this study could yield slightly different conclusions about disparities in evidence-based and non-evidence-based sentencing systems.

Omitted variable bias may be another concern. As noted in Footnote 20, this study's three measures of plea bargaining, offense severity reductions, count reductions, and mandatory minimum reductions, are an imperfect representation of guilty pleas. They do not account for the full range of plea negotiation mechanisms (e.g. sentence and fact bargaining cannot be captured), they do not capture non-negotiated pleas, and they may well include reductions that were the result of something other than a plea deal, such as insufficient evidence. The study also has no reliable way to measure pretrial detention. There is some evidence that sociodemographic disparities characterize the decision to detain offenders pretrial (Kutateladze et al., 2014; Wooldredge et al., 2015) and that pretrial detention influences sentencing outcomes (Lee, 2016; Sacks & Ackerman, 2014), so this pathway for disparities in sentencing cannot be captured. Though the study is able to measure a wide variety of risk factors, there are some additional risk factors that would augment the evaluation of sentencing disparities even more. Gang membership would be especially useful for evaluating racial/ethnic disparities in sentencing. Parenthood, often captured in sentencing research as number of dependents (Koons-Witt, 2002), may be one missing source of gender disparities. Indicators of health and fitness other than alcohol/drug dependency may explain some age disparities. Lastly, though government assistance and perceived financial situation are captured in the LSI-R, a more direct measure of income would help better identify SES disparities. These variables are not available in the dataset.

The study would have further been informative if it had been able to consider disparities in evidence-based and non-evidence-based sentencing for Asian and American offenders more fully. Recent developments in the literature on race and ethnicity emphasize the value of evaluating disparities for minority groups other than blacks (Doerner & Demuth, 2010; Franklin, 2013; Johnson & Betsinger, 2009), and though this study's datasets included cases involving black, Hispanic, Asian, and Native American offenders, only black and Hispanic offenders appeared often enough to ensure the statistical power needed to detect significant disparities. Native Americans appeared so infrequently that the researcher decided to remove them from the sample altogether, and though Asians were included in the analyses, they were likely not represented well enough for the study to make strong conclusions about their treatment in the LSI-R and in the sentencing process.

This study is also unable to distinguish between jail and prison sentences. Prior work indicates that legal and extralegal factors may influence the decision to send an offender to jail and the decision to send an offender to prison differently (Holleran & Spohn, 2004), so the disparities in Connecticut's current system observed in this study may be covering up subtle differences between disparities in jail assignment and disparities in prison assignment. Moreover, by focusing on incarceration and incarceration lengths, this project ignores a variety of other sentencing outcomes (e.g. probation length, fines/restitution, community service, etc.) that the LSI-R and risk factors included in the LSI-R would likely influence in an evidence-based sentencing system. The narrow focus allows this study to zero in on disparities in incarceration, which is consistent with previous literature's attention to the use of jail and prison as punishment during the era of mass incarceration (Cullen et al., 2011). Still, it is crucial to consider and model the entire battery of sentencing options if the field is to understand the full range of sociodemographic disparities in punishment and effects of incorporating risk assessments into sentencing.

One additional remark about punishment measurement is necessary. This study was unable to distinguish between sentences of incarceration and sentences of time served in the available datasets, which resulted in some degree of overestimation of incarceration sentences in the state of Connecticut. This slight distortion of the outcome variable may have an impact on the estimation of disparities in Connecticut's current sentencing scheme, but it became more salient in the context of the evidence-based simulation process, where the researcher found that a nontrivial number of actual incarceration sentences had very low probabilities of evidence-based incarceration. It is likely that in actuality, some of those low-probability cases involved sentences of time served, such that the offenders' experienced punishment differed from what was recorded in the data. Though this fact is unlikely to have drastically altered observed disparities in actual sentences and simulated evidence-based sentences, the limitation should be acknowledged.

The data for this study comes from a single state. Connecticut is arguably an excellent location for a study of the potential impacts of evidence-based sentencing on disparities; its criminal court system has not been subjected to much prior empirical scrutiny, it has broad offense classes with associated penalties rather than sentencing guidelines, it has extensive familiarity with actuarial risk assessments in the criminal justice system, and it has already demonstrated some interest in evidence-based sentencing, which increases the potential utility of the study. But some of these same qualities that make Connecticut an ideal site for the evaluation also make it unique from other states, which limits the generalizability of the study to other locations in the U.S. Findings from this study can only be applied to other states with great caution.

While the study's sampling frame covers the entire state of Connecticut, the only contextual factor that it is able to account for (with clustering) is the court in which each sentence was issued. It does not account for variation in courtroom actors or other courtroom contexts within the state, nor is it able to measure the relationship between any contextual-level characteristics and punishment. Previous literature makes it clear that community, courtroom, and court actor characteristics can have noteworthy effects on criminal justice processing outcomes (Eisenstein & Jacob, 1977; Dixon, 1995; Johnson, 2005). Research on court/community contexts and sentencing finds that a variety of factors such as courtroom caseload, court organizational culture, political environment, unemployment rate, crime rate, and racial and ethnic representation all impact punishment (Britt, 2000; Crow & Johnson, 2008; Ulmer & Johnson, 2004; Wang & Mears, 2015).

Judge-level characteristics may be a particularly influential source of contextual variation in disparities that this study is unable to capture. In Connecticut, where guidelines have not been established and sentencing ranges for various offense classifications are broad, there is abundant room for judicial discretion to enter the punishment decision-making calculus. An evidence-based sentencing scheme likewise relies on judges using their discretion to incorporate actuarial risk information into sentencing decisions. Judge characteristics that shape how discretion is used and how punishment is determined can therefore have substantial effects on disparities in punishment outcomes both in evidence-based and non-evidence-based sentencing schemes. Traditional demographic characteristics like race and gender appear to have only modest and inconsistent effects on punishment and disparity (Spohn & Cederblom, 1991; Johnson, 2006; Steffensmeier & Hebert, 1999), but other more dynamic characteristics such as political liberalism, punishment orientations, prior legal experience, religion, and proximity to reelection times are identified by criminal justice and political science literatures as unique and impactful determinants of sentencing (Gibson, 1978; Gordon & Huber, 2007; Hogarth, 1971; Huber & Gordon, 2004; Myers, 1988). There may be variation in sentencing disparities according to these types of judicial characteristics. The analyses in this study, aggregated across all judges and all Connecticut courtrooms, mask those potentially important variations.

Altogether, this study has limitations stemming from required assumptions, sample selection, potential omitted variable bias and other data constraints, and an inability to capture contextual characteristics. Despite these limitations, the study makes a unique and distinctive contribution to the sentencing literature and to policymakers and practitioners seeking to understand the potential effects of an evidence-based sentencing scheme on punishment and social inequality.

Future Directions for Research

This study builds on existing scholarship to inform the understanding of disparities in evidence-based and non-evidence-based sentencing schemes. At the same time, it also highlights several areas for future research. First, it is important that future sentencing scholarship continues to explore and test mechanisms through which social disparities may emerge in the absence of evidence-based sentencing. This study found that in Connecticut, social and personal characteristics that may help shape judges' perceptions of offenders' recidivism risk explained small but meaningful portions of observed sociodemographic disparity. The risk variables that this study was able to capture are unique, tapping into characteristics like criminal peer influences, quality of living situation, and attitudes toward crime and the criminal justice system that are hypothesized to affect punishment outcomes but are rarely able to be tested (Frazier, 1979; Irwin, 1985; Robinson et al., 1994; Wooldredge, 2007). Similar analyses that attempt to tap into other aspects of offenders' social and personal characteristics may provide additional explanatory power. Innovative studies have begun making strides in that direction. Recent work by Johnson and King (Johnson & King, 2017; King & Johnson, 2016), for example, found that minority offenders were perceived as having more threatening appearances, while facial characteristics such as physical attractiveness, baby-faced appearances, and facial tattoos were also linked to perceptions of threat and to subsequent incarceration outcomes. This type of work shows great promise for improving scholarly understanding of court actor decision-making processes. Future work should continue to examine these types of factors to determine the extent to which they can help explain sociodemographic disparities and punishment outcomes.

Second, the dataset compilation procedures used in this study call attention to the need for a more nuanced understanding of how recidivism risk changes over time. In this study, a risk assessment could be paired with a corresponding case verdict if the assessment occurred no more than one year before or after the verdict date in the case. However, supplementary analyses in the study (see Footnote 8) indicated that risk scores fluctuated substantially across time, raising questions about how risk assessments can and should be used to make punishment decisions in the case of repeat offending. Scholars have previously pointed out the need to evaluate changes in LSI-R score over time (Lowenkamp & Bechtel, 2007). A study by Vose and colleagues (2009) showed that among probationers and parolees, changes in LSI-R scores over time were noticeable, and that the changes even served as significant predictors of subsequent recidivism. Additional work that evaluates changes over time specifically for repeat offenders, and perhaps explores how changes in LSI-R score over time could be used to inform the sentencing decision, would make an interesting line of research.

The application of behavioral economics principles to explain judicial decision-making is a third important area for future work. With its broad focus on choice under uncertainty, behavioral economics has shown its abundant utility across the social sciences and humanities (Slovic et al., 1977), but its principles are only just beginning to percolate into the realm of criminal justice research. This study's theoretical framework unites prominent sociolegal theorizing with core concepts from the behavioral economics literature to provide a more complete narrative of how disparities may emerge in Connecticut's current sentencing scheme and in an evidence-based sentencing scheme. Other exciting work that use principles from the behavioral economics literature such as framing (Leibovitch, 2017; Rachlinski et al., 2015) and anchoring (Englich, 2005; Englich et al., 2006) to explain trends in sentencing outcomes further demonstrates the promise of incorporating behavioral economics approaches into the study of courtroom decision-making. Placing judicial decision-making into utility modeling frameworks, considering the role of concepts like risk aversion and affective bias in judicial decision-making, and identifying particular heuristics that may bias sentencing decisions are potential developments that may assist criminal justice scholars in more fully understanding sociodemographic disparities in punishment.

Evaluation of the factors that shape judges' perceptions of risk is another example of a particular line of inquiry within the larger frame of judicial decision-making that would constitute an important contribution to the empirical and theoretical literature. Despite sociolegal theorizing that posits a link between risk evaluation, demographic characteristics, and punishment, relatively little is known about which offense and offender characteristics judges use to form their impressions of offender risk (Albonetti, 1991; Steffensmeier et al., 1998). Research that explores previously unconsidered mechanisms for shaping judicial perceptions of risk would inform both theory and practice. It is also unclear whether the factors captured in prominent actuarial risk assessments like the LSI-R are the same factors that play into judicial perceptions of risk. This study assumes that judges consider similar social and personal offender characteristics and tap into the same latent risk attribute that the LSI-R does, but it would be interesting and informative to evaluate whether this assumption can be demonstrated empirically. Thoroughly addressing these research questions will likely require the use of qualitative or mixed methods. Though used relatively infrequently in criminal justice research, interview and survey techniques have the potential to provide valuable new insights into decision-making and impression formation processes among sentencing judges and other court actors (e.g. see Kramer & Ulmer, 2009).

The headlining finding in this study, that evidence-based sentencing has large potential to increase social inequality in punishment if implemented in Connecticut, represents an important step forward in the pursuit of understanding the full range of effects of evidence-based sentencing. That said, the study's use of simulation rather than direct observation of evidence-based sentences also underscores the need for research that looks into disparities and other punishment trends in observable evidence-based sentencing schemes. This research would inform the efficacy of this study's simulation approach to policy evaluation as well as provide more nuance about how incorporating actuarial risk assessment into sentencing changes decision-making.

Existing scholarship on evidence-based sentencing is largely non-empirical, focusing on legal and philosophical reasons why actuarial risk prediction and evidence-based sentencing represent either great progress or gross injustice. To be clear, this scholarship makes a substantial contribution to the sentencing literature and has advanced the field's conceptual understanding of risk logics, compatible and incompatible punishment goals, and the constitutionality of risk assessment at sentencing (Hannah-Moffat, 1999, 2013; Monahan & Skeem, 2015; Starr, 2014b). However, it is time for the evidence-based sentencing debates to be informed by data. As evidence-based sentencing becomes more popular among local and state jurisdictions, more and more opportunities for evaluation of sentencing outcomes in evidence-based settings should arise. It is important that researchers capitalize on these opportunities and provide further empirical analysis that can help scholars, policymakers, and practitioners understand and make better, data-driven decisions about sentencing policy.

Ultimately, it will be important to expand this study's inquiries and test this study's research questions in other jurisdictions in the U.S. and in other contexts as well. Connecticut is a single state, and though it may be particularly well-suited for an evaluation of evidence-based sentencing, it has its own unique political, social, and criminal justice cultures that may affect conclusions about social inequality and evidence-based sentencing. Evidence-based sentencing is growing in popularity, and its expansion across the country invites investigation into how it may be used in different ways and may influence social inequality differently in different locations. The practice may also be differentially impactful for sentencing decisions involving specific offense types as well. Though the sample in this study includes both felony and misdemeanor offenses, it is possible that risk plays a different role in determining punishment for more or less serious offenses. For reasons such as these, future research should consider the effects of evidence-based sentencing in jurisdictions outside Connecticut and seek to understand other contextual differences in whether and how punishment disparities emerge in evidence-based and non-evidence-based sentencing schemes.

Policy Implications

One of the strengths of this study is that it has clear and direct policy implications. First, the study provides a snapshot of sentencing outcomes in Connecticut and details the extent to which structural factors (i.e. legal and case processing factors), risk characteristics, and possibly judicial discretion foster sociodemographic disparities in punishment. Results demonstrate that there are indeed significant disparities by race, ethnicity, gender, and SES in punishment outcomes in the state, but that these disparities come from multiple sources. Legal factors, case processing factors, and risk factors all accounted for their own portions of observed disparities. But disparities persisted even after all of these factors were accounted for, suggesting that other aspects of judicial decision-making such as the use of stereotyping and other heuristics may also be contributing to disadvantages for certain offender groups. These findings indicate that tackling sociodemographic disparities in punishment will require a broad intervention strategy, one that acknowledges multiple sources of disadvantage. It should entail the recruitment of multiple criminal justice actors, from police who make discretionary decisions about whom to arrest, to prosecutors who make decisions about who/what to charge, to judges and correctional officers who make punishment and supervision decisions. It should consider the role that both structural sentencing policies and individual discretion play in generating disparate outcomes, offering guidance to policymakers about what types of policy promote fair treatment and to practitioners about how to reduce their own biases. Ultimately, a holistic approach is needed if the criminal justice system is to be made equitable.

Second, the finding that low-SES offenders are substantially disadvantaged across the board in this study has meaningful implications. Offenders with less education, less employment, and poorer financial situations receive harsher sentences in Connecticut's current sentencing scheme, higher LSI-R composite and domain scores, and harsher sentences in the simulated evidence-based sentencing scenario. This information should prompt practitioners and policymakers to discuss why disparities based on SES are so ingrained in the criminal justice system and to think about potential solutions that could change this pattern. There are a variety of directions for these conversations. Discussions may lead to the conclusion that some of the items captured in LSI-R risk assessments are unfit to be determinants of sentencing due to their discriminatory nature. They may also lead to the identification of more treatment and other intervention programs that target the various needs of low-SES offenders specifically. They may even lead to the consideration of adjustments to criminal case processing that would provide low-SES offenders with more effective representation and better equip them to combat unfavorable risk evaluations.

Third, analysis of sociodemographic differences in LSI-R composite and domain scores provides valuable information about areas of greatest need for different offenders. Hannah-Moffat and colleagues argue that actuarial risk assessments like the LSI-R are built on middle-class, white, male norms (Hannah-Moffat, 2009; Hannah-Moffat & Shaw, 2001), and along those same lines, particular risk factors may be more or less prominent for middle-class, white, and male offenders than for other groups. Composite scores showed little variation across racial/ethnic and gender groups, but inspection of individual domain differences revealed that the composite differences were masking far more nuanced patterns of disparity. Black and Hispanic offenders were most disadvantaged in the Criminal History and Education/Employment domains, trends which are consistent with research indicating that minority offenders tend to have longer criminal histories and be less well-off socially and financially than whites (Kutateladze et al., 2015; Spohn, 2000). The Education/Employment disadvantage in particular highlights the linkage between race/ethnicity and SES and should signal to policymakers that continued attention to how these characteristics can be disassociated is warranted. A second finding from the race/ethnicity comparisons, that the Alcohol/Drug and Emotional/Personal domains were a prominent source of relative risk for white offenders, reinforces the need for criminal justice policymakers to identify effective ways to contend with alcohol/drug use and mental illness among offending populations. Together, these differences further hint at the notion that the impetus for offending may differ across racial and ethnic groups, and that effective interventions should take this into account: racial and ethnic minority offenders may be better served by receiving education and employment assistance, while focusing on alcohol/drug treatment programs and mental health services may be more productive for white offenders (Chenane et al., 2015). Similar distinctions can be made for LSI-R domains across gender groups. Consistent with prior literature (Manchak et al., 2009; Mihailides et al., 2005), men in this study scored higher in the Criminal History, Education/Employment, Leisure/Recreation, Companions, Alcohol/Drug, and Attitudes/Orientations domains. Women scored higher in the Emotional/Personal, Financial, Family/Marital, and Accommodations domains. Male-specific interventions, then, may choose to focus on services that assist offenders in areas like job hunting, structuring free time and associating with prosocial peers, and alleviating alcohol/drug dependence. Female offenders, on the other hand, may be more responsive to services that target mental health, financial and residential stability, and family/romantic relationships.

Fourth, this study's analytic strategy shows the utility of prospective simulation techniques for understanding the consequences of future sentencing policies. In this study, evidence-based sentences are a hypothetical rather than a reality, so they had to be simulated rather than observed. Though using a simulation method rather than an observational method ignores the role of judicial discretion in determining punishment and introduces the potential for substantial error in the estimation of each evidence-based sentence, it also allows for a preliminary evaluation of evidence-based sentencing prior to implementation rather than after it. In the case of this study, the benefit of such an approach is that the Connecticut Sentencing Commission, along with Connecticut General Assembly, has the opportunity to make a more informed decision about whether or not to transition to an evidence-based sentencing scheme. Most evaluations of disparities related to criminal justice policies are conducted after the policy has been enacted, such that any negative consequences observed in the evaluation have already borne out in reality (e.g. Crawford et al., 1998; Miethe & Moore, 1985; Ulmer et al., 2007). While these post-implementation evaluations are undoubtedly of great importance, it is also beneficial to consider the potential effects of policies before they are implemented, so that in some cases, the negative consequences can be avoided or at least carefully monitored during initial implementation. Connecticut's own racial impact statement policy embodies this principle; when a racial impact statement is prepared in response to proposed sentencing policies, policymakers can use it to make a more informed, data-driven decision (London, 2011).

The final policy implication to be drawn from this study is that evidence-based sentencing is likely to come at a social cost. Critics of evidence-based sentencing argue that the innovation would worsen social inequality in punishment (Hannah-Moffat, 2013; Monahan & Skeem, 2015; Starr, 2014b,c), and this study provides some indirect evidence in support of that claim. Though disparities were apparent in Connecticut's current sentencing scheme, disparities in a simulated evidence-based scenario were larger in the decision to incarcerate (though not in incarceration lengths). Disparities in scores on actuarial risk assessments like the LSI-R can translate into incarceration disadvantages for groups like racial/ethnic minorities and low-SES offenders. It is important that policymakers considering the implementation of an evidence-based sentencing scheme be aware of this social cost and weight it appropriately against other benefits and costs associated with the practice.

In conclusion, this study informs the contentious debate about the merits and demerits of evidence-based sentencing but also highlights several unanswered questions about how disparities emerge in both evidence-based and non-evidence-based sentencing schemes. Sociodemographic disparities were observed in Connecticut's current sentencing scheme, in LSI-R scores, and in simulated evidence-based sentences derived from those LSI-R scores. Evidence-based sentencing, the results also suggest, has the potential to increase social inequality through its impact on decisions about which offenders to incarcerate. These findings have meaningful implications for scholars interested in understanding disparities in judicial decision-making as well as policymakers seeking to make maximally informed decisions about the future of sentencing policy. The researcher hopes that in the future, more research will focus on the critical issues discussed in this study and further advance the field's understanding of risk evaluation and disparity so that criminal sentencing may become a more data-driven, effective, and equitable enterprise.

TABLES

Table 4.1: Offense Classifications in the Connecticut Penal Code (Title 53a of General Statutes)

Offense Type	Offense Severity	Statutory Sentence
Felony	Capital	Life imprisonment without possibility of parole
	Class A (Murder)	25-60 years imprisonment and/or fines up to $20,000
	Class A	10-25 years imprisonment and/or fines up to $20,000
	Class B	1-20 years imprisonment and/or fines up to $15,000
	Class C	1-10 years imprisonment and/or fines up to $10,000
	Class D	1-5 years imprisonment and/or fines up to $5,000
	Class E	1-3 years imprisonment and/or fines up to $3,500
	Unclassified	Variable
Misdemeanor	Class A	0-1 year imprisonment and/or fines up to $2,000
	Class B	0-6 months imprisonment and/or fines up to $1,000
	Class C	0-3 months imprisonment and/or fines up to $500
	Class D	0-1 month imprisonment and/or fines up to $250
	Unclassified	Variable

Table 7.1: Data Match Rates

Existing Dataset	Added Dataset	Cases From Case Data Successfully Linked	Observations from Added Dataset Successfully Linked
Case Data	Statute Data	103,117 cases (100%)	472 out of 3,933 statutes (12%)
Case Data	Demographic Data	103,117 cases (100%)	63,521 out of 70,220 offenders (90%)
Case Data	LSI-R Data	67,526 out of 103,117 cases (65%)	51,348 out of 115,276 assessments (45%)

Table 7.2: Descriptive Statistics

	Mean	SD	Min	Max
Dependent Variable				
Incarceration (Y/N)	.34	.47	0	1
Incarceration Length in Months	19.22	44.18	.1	720
Independent Variables				
White	.50	.50	0	1
Black	.30	.46	0	1
Hispanic	.20	.40	0	1
Asian	.004	.06	0	1
Female	.17	.38	0	1
Education				
Completed Less Than 10th Grade	.16	.37	0	1
Completed 10th Grade	.38	.48	0	1
High School Degree	.30	.46	0	1
Completed At Least Some College	.16	.36	0	1
Post-Graduate Education	.01	.08	0	1
Employment				
Unemployed, Non-Student	.62	.49	0	1
Part-Time Employed or Student	.10	.30	0	1
Full-Time Employed or Student	.28	.45	0	1
Financials				
Reliant Upon Social Assistance	.49	.50	0	1
Financial Problems	1.96	.74	0	3
Legal and Case Processing Factors				
Offense Type				
Violent	.15	.35	0	1
Property	.22	.42	0	1
Drug	.30	.46	0	1
Sex	.01	.12	0	1
Weapons	.02	.14	0	1
Public Order	.30	.46	0	1
Offense Severity Class				
Class A Murder	.000	.01	0	1
Class A Felony	.002	.04	0	1
Class B Felony	.03	.17	0	1
Class C Felony	.05	.23	0	1
Class D Felony	.14	.35	0	1
Class E Felony	.005	.07	0	1
Unclassified Felony	.15	.36	0	1
Class A Misdemeanor	.34	.48	0	1
Class B Misdemeanor	.16	.37	0	1
Class C Misdemeanor	.10	.30	0	1
Class D Misdemeanor	.005	.07	0	1
Unclassified Misdemeanor	.000	.01	0	1
Criminal History				

Prior Adult Convictions	3.28	3.48	0	12
LSI-R Criminal History Score	4.73	2.45	0	10
Number of Current Convictions	1.68	1.03	1	5
Under Probation Supervision	.15	.36	0	1
Mandatory Minimum Applied	.01	.11	0	1
Counts Dropped After Charging	.57	.50	0	1
Severity Reduced After Charging	.28	.45	0	1
MM Dropped After Charging	.02	.15	0	1
Conviction Year				
FY2008	.34	.47	0	1
FY2009	.34	.48	0	1
FY2010	.32	.47	0	1
Risk Factors				
Age				
18-24	.34	.47	0	1
25-34	.28	.45	0	1
35-44	.21	.41	0	1
45+	.17	.38	0	1
Married	.10	.30	0	1
Family/Marital Domain Score	2.04	1.18	0	4
Accommodation Domain Score	.99	.94	0	3
Leisure/Recreation Domain Score	1.60	.65	0	2
Companions Domain Score	2.74	1.29	0	5
Alcohol/Drug Domain Score	4.26	2.56	0	9
Emotional/Personal Domain Score	1.70	1.55	0	5
Attitudes/Orientations Domain Score	1.37	1.29	0	4

N=66,984

Table 7.3: Associations Among SES Components (Goodman & Kruskal's Gamma Coefficients)[50]

	Education	Employment	Financial: Social Assistance	Financial: Problems (0-3)
Education (0-4)	1.00			
Employment (0-2)	.16	1.00		
Financial: Social Assistance	-.15	-.55	1.00	
Financial: Problems (0-3)	-.09	-.53	.47	1.00

[50] Education and employment are both measured using series of dichotomous variables in all analyses. However, the education/employment categories can easily be combined into meaningful ordinal variables, in which higher numbers represent more education/employment. The ordinal education variable contains five categories ranging from less than ninth grade completed to post-grad education, while the ordinal employment variable contains three categories ranging from unemployed to full-time employed/student. In order to capture the full association between SES components, the researcher uses these ordinal variables rather than using each dichotomous variable individually.

Table 7.4: Descriptive Statistics Comparing Matched and Unmatched Case Data

	Unmatched or Incomplete Cases (N=36,133)	Matched Cases/Final Sample (N=66,984)
	Mean	Mean
Dependent Variable		
Incarceration (Y/N)	.46	.34
Incarceration Length in Months	25.91	19.22
Independent Variables		
White	.40	.50
Black	.36	.30
Hispanic	.23	.20
Asian	.00	.004
Female	.15	.17
Legal and Case Processing Factors		
Offense Type		
Violent	.12	.15
Property	.22	.22
Drug	.26	.30
Sex	.02	.01
Weapons	.02	.02
Public Order	.36	.30
Offense Severity Class		
Class A Murder	.000	.000
Class A Felony	.003	.002
Class B Felony	.03	.03
Class C Felony	.07	.05
Class D Felony	.13	.14
Class E Felony	.004	.005
Unclassified Felony	.14	.15
Class A Misdemeanor	.32	.34
Class B Misdemeanor	.15	.16
Class C Misdemeanor	.15	.10
Class D Misdemeanor	.005	.005
Unclassified Misdemeanor	.000	.000
Prior Adult Convictions	4.04	3.28
Number of Current Convictions	1.57	1.68
Under Probation Supervision	.11	.15
Mandatory Minimum Applied	.02	.01
Counts Dropped After Charging	.55	.57
Severity Reduced After Charging	.27	.28
MM Dropped After Charging	.02	.02
Conviction Year		
FY2008	.37	.34
FY2009	.33	.34
FY2010	.30	.32

Age		
18-24	.26	.34
25-34	.30	.28
35-44	.25	.21
45+	.19	.17

Table 8.1: Aggregate Disparities in Punishment

	Decision to Incarcerate		Sentence Length		
	Odds Ratio		Length	AME	
Constant	.24 (.02)	***	-11.31 (.19)	--	***
Independent Variables					
Black	1.09 (.06)		.56 (.13)	8.25	***
Hispanic	.94 (.04)		.42 (.15)	6.25	**
Asian	.99 (.14)		.49 (.59)	6.91	
Male	2.17 (.07)	***	.49 (.06)	7.21	***
Completed 10th Grade	.99 (.03)		-.13 (.06)	-1.99	*
High School Degree	.86 (.03)	***	-.36 (.07)	-5.38	***
Completed At Least Some College	.88 (.04)	**	-.44 (.07)	-6.33	***
Post-Graduate Education	.98 (.13)		-.51 (.21)	-7.56	*
Part-Time Employed or Student	.50 (.02)	***	-.50 (.11)	-7.38	***
Full-Time Employed or Student	.47 (.02)	***	-.49 (.07)	-7.40	***
Reliant Upon Social Assistance	.90 (.04)	*	-.38 (.08)	-5.78	***
Financial Problems	1.17 (.02)	***	-.01 (.05)	-.03	
McFadden's Pseudo- R^2	.04		.01		

N=66,984
Models are estimated using a negative binomial hurdle model
Standard errors in parentheses, clustered by court (n=37)
* $p<.05$; ** $p<.01$; *** $p<.001$
Average marginal effect (AME) measured in months

Table 8.2: Disparities in Punishment, Accounting for Legal Factors

	Decision to Incarcerate		Sentence Length		
	Odds Ratio		Length	AME	
Constant	.04 (.01)	***	.59 (.09)	--	***
Independent Variables					
Black	.86 (.06)	***	.24 (.04)	4.06	***
Hispanic	.87 (.04)	**	.18 (.05)	3.02	***
Asian	1.08 (.16)		-.17 (.17)	-3.33	
Male	1.60 (.05)	***	.15 (.04)	2.56	***
Completed 10th Grade	.98 (.03)		-.00 (.04)	.12	
High School Degree	.96 (.04)		-.09 (.03)	-1.54	**
Completed At Least Some College	1.08 (.05)	*	-.12 (.04)	-1.97	**
Post-Graduate Education	1.49 (.22)	**	-.01 (.19)	-.44	
Part-Time Employed or Student	.61 (.03)	***	-.07 (.04)	-1.20	
Full-Time Employed or Student	.58 (.02)	***	-.11 (.04)	-1.87	**
Reliant Upon Social Assistance	.82 (.03)	***	-.09 (.03)	-1.53	**
Financial Problems	1.09 (.02)	***	.01 (.02)	.15	
Legal and Case Processing Factors					
Violent	1.35 (.07)	***	.39 (.04)	7.25	***
Property	1.01 (.05)		.29 (.04)	4.97	***
Drug	1.99 (.10)	***	-.74 (.09)	-9.04	***
Sex	1.94 (.26)	***	.48 (.06)	8.28	***
Weapons	2.02 (.36)	***	.42 (.06)	7.59	***
Class A Murder	--		3.71 (.20)	70.01	***
Class A Felony	85.73 (89.53)	***	2.98 (.11)	55.97	***

Class B Felony	14.04 (2.02)	***	2.11 (.06)	38.98	***
Class C Felony	4.98 (.35)	***	1.27 (.06)	23.06	***
Class D Felony	2.81 (.17)	***	.92 (.04)	16.63	***
Class E Felony	1.08 (.19)		.16 (.14)	2.98	
Unclassified Felony	2.05 (.27)	***	2.00 (.09)	31.70	***
Class B Misdemeanor	.73 (.03)	***	-.86 (.10)	-10.37	***
Class C Misdemeanor	.87 (.08)		-.68 (.12)	-9.80	***
Class D Misdemeanor	.71 (.15)		-.98 (.42)	-14.31	*
Unclassified Misdemeanor	.50 (.48)		-25.88 (.10)	-34.87	***
Prior Adult Convictions	1.03 (.01)	**	.03 (.01)	.46	***
LSI-R Criminal History Score	1.19 (.01)	***	.02 (.01)	.43	***
Number of Current Convictions	1.41 (.03)	***	.24 (.02)	4.34	***
Under Probation Supervision	1.83 (.15)	***	.28 (.06)	4.23	***
Mandatory Minimum Applied	3.48 (.74)	***	.54 (.04)	10.92	***
Counts Dropped After Charging	.88 (.06)		-.32 (.04)	-5.34	***
Severity Reduced After Charging	.95 (.03)		.38 (.04)	6.11	***
MM Dropped After Charging	1.18 (.10)	*	.23 (.04)	4.57	***
FY2009	1.08 (.05)		.03 (.03)	.52	
FY2010	1.06 (.06)		.06 (.03)	1.05	*
McFadden's Pseudo- R^2	.20		.06		

N=66,984
Models are estimated using a negative binomial hurdle model
Standard errors in parentheses, clustered by court (n=37)
* p<.05 ; ** p<.01 ; *** p<.001
Average marginal effect (AME) measured in months

Table 8.3: Disparities in Punishment, Accounting for Legal and Risk Factors

	Decision to Incarcerate		Sentence Length		
	Odds Ratio		Length	AME	
Constant	.03 (.01)	***	.34 (.11)	--	**
Independent Variables					
Black	.80 (.05)	***	.21 (.03)	3.28	***
Hispanic	.83 (.04)	***	.16 (.04)	2.47	***
Asian	1.08 (.16)		-.12 (.14)	-2.43	
Male	1.51 (.06)	***	.16 (.04)	2.50	***
Completed 10th Grade	.98 (.03)		.00 (.03)	.07	
High School Degree	.99 (.04)		-.06 (.02)	-1.05	**
Completed At Least Some College	1.11 (.05)	**	-.07 (.04)	-1.03	*
Post-Graduate Education	1.38 (.20)	*	.02 (.16)	.30	
Part-Time Employed or Student	.63 (.03)	***	-.07 (.04)	-1.13	
Full-Time Employed or Student	.61 (.02)	***	-.09 (.03)	-1.42	**
Reliant Upon Social Assistance	.82 (.02)	***	-.10 (.02)	-1.49	***
Financial Problems	1.04 (.02)		-.01 (.02)	-.18	
Legal and Case Processing Factors					
Violent	1.43 (.08)	***	.36 (.03)	6.13	***
Property	1.00 (.05)		.27 (.04)	4.28	***
Drug	2.48 (.14)	***	-.79 (.09)	-8.18	***
Sex	2.01 (.27)	***	.48 (.06)	7.53	***
Weapons	2.04 (.37)	***	.32 (.06)	5.26	***
Class A Murder	--		3.71 (.15)	64.41	***
Class A Felony	92.45 (96.38)	***	3.04 (.11)	52.42	***

Class B Felony	15.01 (2.19)	***	2.18 (.06)	39.97	***
Class C Felony	5.34 (.38)	***	1.32 (.06)	21.90	***
Class D Felony	2.95 (.17)	***	.99 (.04)	16.26	***
Class E Felony	1.15 (.20)		.32 (.14)	5.41	*
Unclassified Felony	1.69 (.23)	***	2.14 (.10)	30.40	***
Class B Misdemeanor	.99 (.05)		-.85 (.10)	-8.73	***
Class C Misdemeanor	.88 (.08)		-.60 (.13)	-7.67	***
Class D Misdemeanor	.81 (.18)		-.87 (.39)	-11.69	*
Unclassified Misdemeanor	.50 (.46)		-1.02 (.40)	-26.97	*
Prior Adult Convictions	1.01 (.01)		.03 (.01)	.51	***
LSI-R Criminal History Score	1.19 (.01)	***	.01 (.01)	.16	
Number of Current Convictions	1.40 (.03)	***	.17 (.01)	2.89	***
Under Probation Supervision	1.72 (.13)	***	.36 (.06)	5.20	***
Mandatory Minimum Applied	3.58 (.76)	***	.53 (.03)	9.83	***
Counts Dropped After Charging	.92 (.06)		-.29 (.04)	-4.34	***
Severity Reduced After Charging	.92 (.03)	*	.35 (.04)	4.96	***
MM Dropped After Charging	1.22 (.10)	*	.20 (.03)	3.69	***
FY2009	1.04 (.05)		.03 (.03)	.42	
FY2010	1.01 (.06)		.06 (.03)	1.03	*
Risk Factors			.07 (.03)		
Age 25-34	1.00 (.03)		.02 (.03)	.32	
Age 35-44	1.02 (.04)		-.04 (.03)	-.52	
Age 45+	1.12 (.05)	**	-.17 (.04)	-2.39	***
Married	1.01 (.05)		.05 (.04)	.71	**
Repeat Conviction in Study Period	1.14 (.02)	***	-.02 (.02)	.29	

Family/Marital Domain Score	.97 (.01)	*	.01 (.01)	.17	**
Accommodation Domain Score	1.14 (.02)	***	.02 (.01)	.34	
Leisure/Recreation Domain Score	1.04 (.03)		-.04 (.02)	-.57	*
Companions Domain Score	.99 (.02)		.07 (.01)	1.10	***
Alcohol/Drug Domain Score	1.02 (.01)	*	-.01 (.01)	-.19	*
Emotional/Personal Domain Score	.95 (.01)	***	.02 (.01)	.23	*
Attitudes/Orientations Domain Score	1.02 (.02)		.03 (.01)	.50	***
McFadden's Pseudo- R^2	.21		.07		

N=66,984
Models are estimated using a negative binomial hurdle model
Standard errors in parentheses, clustered by court (n=37)
* p<.05 ; ** p<.01 ; *** p<.001
Average marginal effect (AME) measured in months

Table 8.4: Aggregate LSI-R Composite and Domain Score Differences by Race, Ethnicity, Gender, and SES

	Composite Differences	Domain Score Differences									
		CH	EE	FI	FM	AC	LE	CO	AD	EP	AO
Black	1.17	.73	.96	.08	.04	.36	.10	.37	-1.05	-.69	.27
Hispanic	.65	.25	1.13	.09	**.01**	.32	.13	.21	-.98	-.51	**.03**
Asian	-5.05	-1.26	-.70	-.28	-.55	**-.10**	**.04**	.26	-1.38	-.58	**.03**
Male	**-.02**	.87	.11	-.28	-.34	-.05	.07	.11	.14	-.84	.20
Completed 10th Grade	-1.10	.23	-1.04	-.06	**-.04**	-.14	-.07	**-.02**	.11	**-.06**	**-.02**
High School Degree	-4.33	-.36	-2.80	-.12	-.24	-.30	-.12	-.33	.11	**-.05**	-.12
At Least Some College	-6.36	-.77	-3.44	-.22	-.44	-.46	-.31	-.70	**.08**	.13	-.24
Post-Graduate Education	-11.73	-2.15	-4.91	-.57	-.78	-.68	-.62	-1.45	-.58	.45	-.44
Part-Time Employed/ Student	-3.64	-1.16	**-.07**	-.35	-.19	-.28	-.11	-.30	-.71	-.33	-.16
Full-Time Employed or Student	-10.84	-1.28	-3.88	-.68	-.42	-.48	-.31	-.54	-.86	-.46	-.25
Reliant Upon Social Assistance	6.27	1.04	1.22	1.24	.43	.45	.13	.40	.59	.65	.13
Relatively Satisfactory Finances	3.81	.93	.89	.10	.28	.14	.27	.49	.57	**-.02**	.17
Relatively Unsatisfactory Finances	10.48	1.90	2.39	1.35	.78	.56	.55	.90	1.36	.25	.44
Very Unsatisfactory Finances	13.90	2.26	3.28	1.51	1.14	.94	.64	1.23	1.76	.49	.67

N=66,984

CH=Criminal History ; EE=Education/Employment ; FI=Finances ; FM=Family/Marital ; AC=Accommodations ; LE=Leisure ; CO=Companions ; AD=Alcohol/Drug ; EP=Emotional/Personal ; AO=Attitudes/Orientations

Reference groups for categorical outcomes are: white, female, completed <10th grade, unemployed, and very satisfactory finances.

Values in the table indicate the score difference between the specified group and the reference group

Kruskal-Wallis tests are used to evaluate statistical significance.

Bolded values are NOT significant at p<.001

Table 8.5: LSI-R Composite and Domain Score Differences by Race, Ethnicity, Gender, and SES

	Composite*	CH	EE	FM	AC	LE	CO	AD	EP	AO
Black	-.32 (.06)	1.50 (.02)	1.42 (.02)	.93 (.02)	1.96 (.03)	1.20 (.02)	1.45 (.02)	.42 (.01)	.40 (.01)	1.35 (.02)
Hispanic	-1.58 (.07)	.95 (.02)	1.05 (.02)	.79 (.02)	1.70 (.03)	1.20 (.03)	**1.03** **(.02)**	.43 (.01)	.51 (.01)	.93 (.02)
Asian	-2.97 (.38)	.48 (.05)	.84 (.09)	.52 (.06)	**1.11** **(.13)**	1.32 (.17)	**.87** **(.09)**	.42 (.05)	.59 (.06)	**1.16** **(.12)**
Male	1.18 (.06)	2.24 (.04)	1.29 (.02)	.64 (.01)	**1.00** **(.02)**	1.32 (.03)	1.22 (.02)	1.31 (.02)	.47 (.01)	1.37 (.03)
Completed 10th Grade	-.82 (.07)	1.26 (.03)	.27 (.01)	.93 (.02)	.84 (.02)	.81 (.02)	**.99** **(.02)**	**1.02** **(.02)**	.92 (.02)	**.97** **(.02)**
High School Degree	-3.63 (.08)	.86 (.02)	.05 (.00)	.68 (.02)	.68 (.02)	.71 (.02)	.66 (.02)	.96 (.02)	.85 (.02)	.87 (.02)
At Least Some College	-4.91 (.09)	.75 (.02)	.03 (.00)	.51 (.02)	.53 (.01)	.46 (.01)	.43 (.01)	.92 (.02)	**.96** **(.02)**	.77 (.02)
Post-Graduate Education	-8.36 (.32)	.36 (.03)	.01 (.00)	.36 (.03)	.43 (.05)	.27 (.03)	.18 (.02)	.60 (.06)	1.38 (.13)	.65 (.06)
Part-Time Employed/ Student	-1.96 (.08)	.51 (.01)	.92 (.02)	.93 (.02)	.79 (.02)	.80 (.02)	.77 (.02)	.70 (.02)	.77 (.02)	.91 (.02)
Full-Time Employed or Student	-5.82 (.06)	.57 (.01)	.03 (.00)	.84 (.01)	.67 (.01)	.53 (.01)	.68 (.01)	.65 (.01)	.74 (.01)	.87 (.02)
Reliant Upon Social Assistance	3.08 (.05)	1.73 (.03)	**1.01** **(.02)**	1.35 (.02)	1.66 (.03)	**.99** **(.02)**	1.30 (.02)	1.27 (.02)	1.75 (.03)	1.03 (.02)
Relatively Satisfactory Finances	2.47 (.20)	1.74 (.10)	1.39 (.08)	1.43 (.08)	1.29 (.09)	1.69 (.10)	1.77 (.10)	1.42 (.08)	**.97** **(.06)**	1.30 (.08)
Relatively Unsatisfactory Finances	6.80 (.20)	2.77 (.15)	2.55 (.15)	2.73 (.16)	2.66 (.18)	3.32 (.20)	2.66 (.16)	2.17 (.12)	**1.08** **(.06)**	1.85 (.11)
Very Unsatisfactory Finances	8.66 (.20)	2.97 (.17)	3.45 (.20)	4.36 (.26)	4.77 (.33)	4.23 (.27)	3.91 (.24)	2.73 (.15)	1.24 (.07)	2.45 (.15)

N=66,984

CH=Criminal History ; EE=Education/Employment ; FM=Family/Marital ; AC=Accommodations ; LE=Leisure ; CO=Companions ; AD=Alcohol/Drug ; EP=Emotional/Personal ; AO=Attitudes/Orientations

Finances domain score is represented in its entirety in the independent variables and is excluded from this analysis
*Composite model is estimated using an ordinary least squares regression
All domain models are estimated using an ordered logistic regression
Coefficients for ordered logistic regressions are odds ratios
Bolded odds ratios NOT significant at $p<.05$

Table 8.6: Negative Binomial Hurdle Model Establishing the Relationship Between Risk and Punishment

	Decision to Incarcerate		Sentence Length		
	OR		Length	AME	
Legal and Case Processing Factors					
Violent	1.38 (.07)	***	.40 (.04)	7.33	***
Property	.99 (.05)		.26 (.04)	4.51	***
Drug	2.58 (.15)	***	-.76 (.09)	9.27	***
Sex	1.94 (.24)	***	.48 (.06)	8.25	***
Weapons	2.08 (.37)	***	.50 (.05)	8.91	***
Class A Murder	--		3.79 (.20)	71.27	***
Class A Felony	93.62 (97.40)	***	3.04 (.10)	56.88	***
Class B Felony	14.99 (2.09)	***	2.15 (.07)	39.64	***
Class C Felony	5.14 (.37)	***	1.29 (.06)	23.35	***
Class D Felony	2.94 (.17)	***	.93 (.04)	16.76	***
Class E Felony	1.08 (.18)		.15 (.14)	2.81	
Unclassified Felony	1.65 (.22)	***	2.06 (.10)	32.73	***
Class B Misdemeanor	1.00 (.04)		-.91 (.10)	11.09	***
Class C Misdemeanor	.90 (.08)		-.71 (.12)	10.27	***
Class D Misdemeanor	.88 (.18)		-.97 (.45)	14.01	*
Unclassified Misdemeanor	.65 (.62)		-18.50 (.13)	35.83	***
Prior Adult Convictions	1.02 (.01)	*	.03 (.01)	.44	***
LSI-R Criminal History Score	1.17 (.01)	***	.02 (.01)	.30	**
Number of Current Convictions	1.42 (.03)	***	.23 (.02)	4.20	***
Under Probation Supervision	1.71 (.14)	***	.30 (.06)	4.56	***
Mandatory Minimum Applied	3.86 (.80)	***	.59 (.04)	11.67	***
Counts Dropped After Charging	.92		-.33	5.56	***

	(.06)		(.04)		
Severity Reduced After Charging	.90 (.03)	**	.41 (.04)	6.53	***
MM Dropped After Charging	1.26 (.10)	**	.24 (.04)	4.74	***
FY2009	1.08 (.05)		.01 (.03)	.26	
FY2010	1.08 (.06)		.04 (.03)	.72	
LSI-R Composite Score	1.02 (.00)	***	.01 (.00)	.18	***
McFadden's Pseudo- R^2	.19		.06		

N=66,984

Standard errors in parentheses, clustered by court (n=37)

* p<.05 ; ** p<.01 ; *** p<.001

Table 8.7: Comparison of Simulated Evidence-Based Sentencing Distributions

	% Incarcerated	Average Sentence Length
Actual Sentences	34	19.22 (44.18)
Simulated Sentences		
.4 Threshold	33	24.36 (33.60)
.5 Threshold	23	30.11 (38.92)
.6 Threshold	15	37.90 (45.79)

N=66,984

Standard errors in parentheses

Table 8.8: Disparities in Simulated Evidence-Based Punishment, Using .4 Threshold

	Decision to Incarcerate		Sentence Length		
	Odds Ratio		Length	AME	
Constant	.20 (.02)	***	6.72 (.13)	--	***
Independent Variables					
Black	1.38 (.07)	***	.11 (.07)	2.63	
Hispanic	1.05 (.05)		.07 (.08)	1.73	
Asian	.78 (.10)		.24 (.14)	5.77	
Male	1.98 (.06)	***	.13 (.04)	3.29	***
Completed 10th Grade	1.00 (.03)		-.11 (.05)	-2.62	*
High School Degree	.78 (.04)	***	-.14 (.05)	-3.53	**
Completed At Least Some College	.68 (.03)	***	-.14 (.05)	-3.53	**
Post-Graduate Education	.48 (.07)	***	-.21 (.08)	-5.17	**
Part-Time Employed or Student	.58 (.02)	***	-.21 (.06)	-5.19	***
Full-Time Employed or Student	.51 (.02)	***	-.19 (.05)	-4.64	***
Reliant Upon Social Assistance	1.19 (.06)	**	-.22 (.06)	-5.37	***
Financial Problems	1.26 (.02)	***	-.01 (.02)	-.32	
McFadden's Pseudo- R^2	.05		.00		

N=66,984
Models are estimated using a negative binomial hurdle model
Standard errors in parentheses, clustered by court (n=37)
* p<.05 ; ** p<.01 ; *** p<.001
Average marginal effect (AME) measured in months

Table 8.9: Disparities in Simulated Evidence-Based Punishment, Using .5 Threshold

	Decision to Incarcerate		Sentence Length		
	Odds Ratio		Length	AME	
Constant	.12 (.02)	***	6.93 (.13)	--	***
Independent Variables					
Black	1.43 (.08)	***	.09 (.07)	2.61	
Hispanic	1.07 (.06)		.06 (.09)	1.96	
Asian	.79 (.12)		.28 (.16)	8.41	
Male	2.04 (.08)	***	.14 (.04)	4.36	***
Completed 10th Grade	1.02 (.03)		-.12 (.05)	-3.67	*
High School Degree	.73 (.03)	***	-.12 (.05)	-3.57	*
Completed At Least Some College	.64 (.04)	***	-.12 (.05)	-3.66	*
Post-Graduate Education	.41 (.08)	***	-.20 (.09)	-6.03	*
Part-Time Employed or Student	.56 (.03)	***	-.21 (.06)	-6.33	***
Full-Time Employed or Student	.50 (.03)	***	-.19 (.05)	-5.65	***
Reliant Upon Social Assistance	1.13 (.07)	*	-.24 (.06)	-7.11	***
Financial Problems	1.23 (.03)	***	-.02 (.03)	-.52	
McFadden's Pseudo- R^2	.05		.00		

N=66,984
Models are estimated using a negative binomial hurdle model
Standard errors in parentheses, clustered by court (n=37)
* p<.05 ; ** p<.01 ; *** p<.001
Average marginal effect (AME) measured in months

Table 8.10: Disparities in Simulated Evidence-Based Punishment, Using .6 Threshold

	Decision to Incarcerate		Sentence Length		
	Odds Ratio		Length	AME	
Constant	.08 (.01)	***	7.14 (.14)	--	***
Independent Variables					
Black	1.42 (.10)	***	.09 (.08)	3.37	
Hispanic	1.01 (.07)		.10 (.10)	3.64	
Asian	.75 (.19)		.33 (.17)	12.45	
Male	2.15 (.09)	***	.16 (.04)	6.00	***
Completed 10th Grade	.97 (.04)		-.12 (.05)	-4.56	*
High School Degree	.70 (.03)	***	-.11 (.05)	-4.31	*
Completed At Least Some College	.61 (.04)	***	-.11 (.05)	-4.26	*
Post-Graduate Education	.34 (.07)	***	-.19 (.12)	-7.24	
Part-Time Employed or Student	.57 (.04)	***	-.25 (.06)	-9.58	***
Full-Time Employed or Student	.50 (.03)	***	-.19 (.05)	-7.39	***
Reliant Upon Social Assistance	1.05 (.06)		-.23 (.06)	-8.87	***
Financial Problems	1.21 (.04)	***	-.02 (.03)	-.65	
McFadden's Pseudo- R^2	..04		.00		

N=66,984
Models are estimated using a negative binomial hurdle model
Standard errors in parentheses, clustered by court (n=37)
* p<.05 ; ** p<.01 ; *** p<.001
Average marginal effect (AME) measured in months

Table 8.11: Comparison of Disparities in Actual and Simulated Evidence-Based Decision to Incarcerate

	Actual Decision To Incarcerate	Evidence-Based Decision to Incarcerate	Direction of Change	
	Odds Ratio	Odds Ratio		
Black	1.09	1.38	↑	***
Hispanic	.94	1.05	Switch	***
Asian	.99	.78	↑	
Male	2.17	1.98	↓	**
Completed 10th Grade	.99	1.00	↓	
High School Degree	.86	.78	↑	***
Completed At Least Some College	.88	.68	↑	***
Post-Graduate Education	.98	.48	↑	***
Part-Time Employed or Student	.50	.58	↓	**
Full-Time Employed or Student	.47	.51	↓	*
Reliant Upon Social Assistance	.90	1.19	Switch	***
Financial Problems	1.17	1.26	↑	***

↑ = evidence-based sentences show more disparity

↓ = evidence-based sentences show less disparity

Switch = disparity switches direction in evidence-based sentences

Evidence-based sentences are drawn from the .4 threshold model.

Significance tests are conducted using coefficients from a seemingly unrelated regressions model.

Table 8.12: Comparison of Disparities in Actual and Simulated Evidence-Based Incarceration Length

	Actual Sentence Length	Evidence-Based Sentence Length	Difference	
	AME	AME		
Black	8.25	2.63	↓	***
Hispanic	6.25	1.73	↓	***
Asian	6.91	5.77	↓	
Male	7.21	3.29	↓	***
Completed 10^{th} Grade	-1.99	-2.62	↑	
High School Degree	-5.38	-3.53	↓	***
Completed At Least Some College	-6.33	-3.53	↓	***
Post-Graduate Education	-7.56	-5.17	↓	
Part-Time Employed or Student	-7.38	-5.19	↓	***
Full-Time Employed or Student	-7.40	-4.64	↓	***
Reliant Upon Social Assistance	-5.78	-5.37	↓	***
Financial Problems	-.03	-.32	↑	

↑ = evidence-based sentence shows more disparity

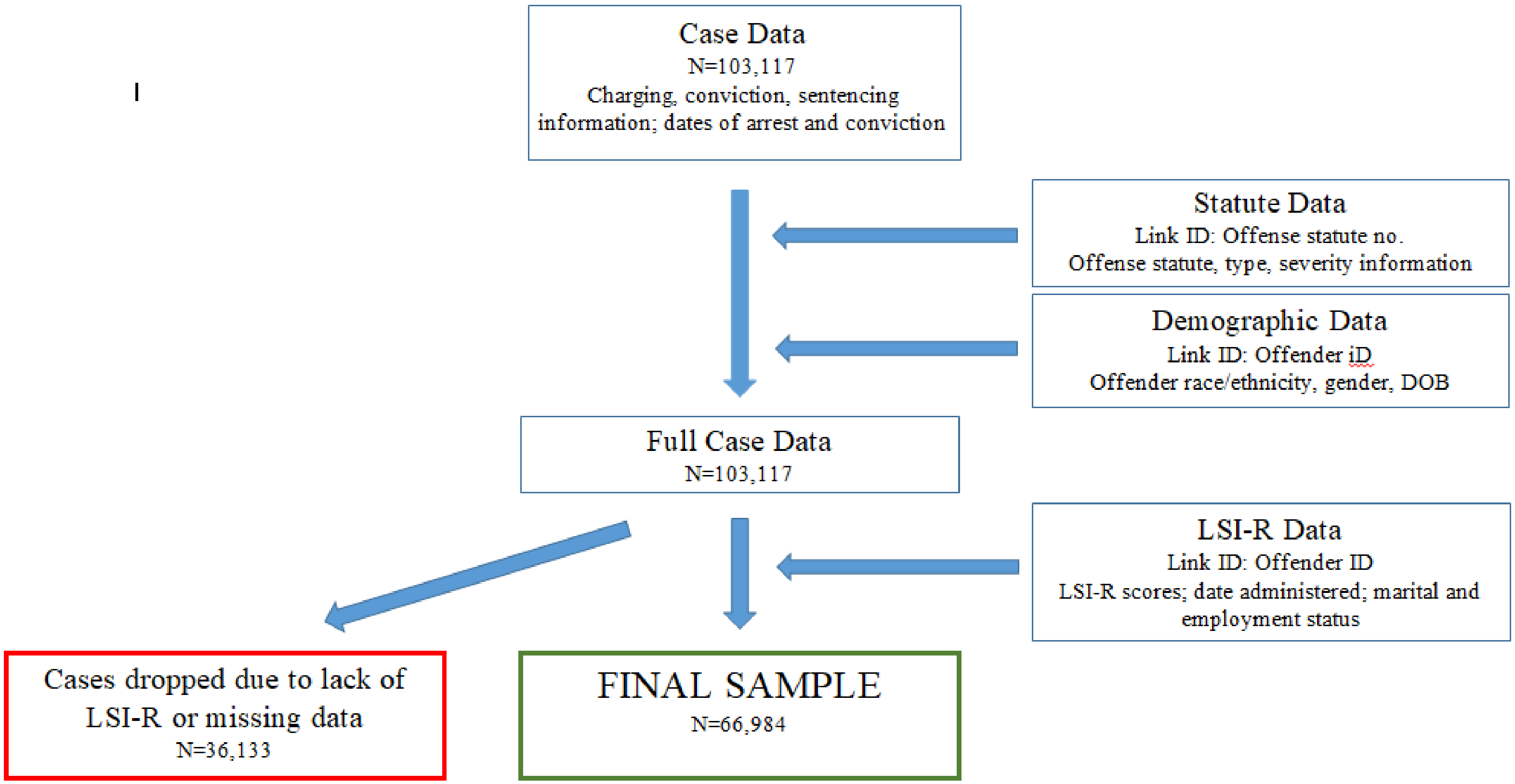
↓ = evidence-based sentence shows less disparity

Evidence-based sentence lengths are drawn from the .4 threshold model.

Significance tests are conducted using coefficients from a seemingly unrelated regressions model.

FIGURES

Figure 6.1: Data Compilation Process



APPENDICES

Appendix A: Mandatory Minimum Sentences and Presumptive Sentences, 2008[51]

Classification	**Statute §**	**Offense**	**Mandatory Minimum Sentence**
Capital Felony	53a-54b	Capital Felony	Execution or life imprisonment without possibility of release
Class A Felony	53a-54a	Murder	25 years
	53a-54c	Felony Murder	25 years
	53a-70c	Aggravated sexual assault of a minor	1st offense: 25 years Subsequent offense: 50 years
	53a-59c	Assault of pregnant woman resulting in termination of pregnancy	10 years
	53a-70	Sexual assault 1st degree (wen use force and victim under 16 or victim under 13 and actor more than two years older)	5 years; 20 years if use force; 10-year mandatory minimum period of combined imprisonment and special parole
	53a-70a	Aggravated sexual assault 1st degree (victim under 16)	5 years; 20 years if use force; 10-year mandatory minimum period of combined imprisonment and special parole
	53a-92	Kidnapping 1st degree	1 year
	53a-92a	Kidnapping 1st degree with a firearm	1 year
	53a-111	Arson 1st degree	None
	53a-196a	Employing a minor in an obscene performance	10 years
	PA 08-1	Home invasion	10 years
	--	All other Class A Felonies	10 years
Class B Felony	53a-55a	Manslaughter 1st degree with a firearm	5 years
	53a-59(a)(1)	Assault 1st degree (when cause injury using deadly weapon or dangerous	5 years; 10 years if victim is under age 10 or a witness

[51] Information in table drawn from Reinhart (2008).

		instrument)	
	53a-59a	Assault of an elderly, blind, disabled, pregnant, or mentally retarded person 1st degree	5 years
	53a-70a	Aggravated sexual assault 1st degree (victim over age 16)	5 years; 10-year mandatory minimum period of combined imprisonment and special parole
	53a-94a	Kidnapping 2nd degree with a firearm	3 years
	53a-101(a)(1)	Burglary 1st degree (when committed with explosive, deadly weapon, or dangerous instrument)	5 years
	53a-134(a)(2)	Robbery 1st degree (when committed with deadly weapon)	5 years
	53a-59	Assault 1st degree (except when use deadly weapon or dangerous instrument)	10 years if victim under age 10 or a witness
	53a-70	Sexual assault 1st degree	2 years; 10 years if victim under age 10; 10-year mandatory minimum period of combined imprisonment and special parole
	53a-71	Sexual assault 2nd degree (victim under 16)	9 months
	53a-72b	Sexual assault 3rd degree with a firearm (victim under 16)	2 years
	53a-90a	Enticing a minor (victim under age 13 or 3rd or subsequent offense with victim 13 or older)	When minor under age 13: 1st offense: 5 years Subsequent offense: 10 years
	53a-94	Kidnapping 2nd degree	3 years
	53a-196c	Importing child pornography	5 years
	53a-196d	Possessing child pornography 1st degree	5 years
	53a-301	Computer crime in furtherance of terrorism	When directed toward public safety agency: 5 years
	53-21	Injury or risk of injury to a minor (involving contact with intimate parts of a minor under age 13)	5 years

Class C Felony	53a-56a	Manslaughter 2nd degree with a firearm	1 year
	53a-71	Sexual assault 2nd degree (victim 16 or older)	9 months
	53a-72b	Sexual assault 3rd degree with a firearm (victim 16 or older)	2 years
	53a-102a	Burglary 2nd degree with a firearm	1 year
	53a-165aa	Hindering prosecution 1st degree	5 years
	53a-196e	Possessing child pornography 2nd degree	2 years
	53a-303	Contamination of public water or food for terrorism	5 years
	53-202b	Selling or transporting assault weapons	2 years; 6 years if sale is to a minor
Class D Felony	53a-216	Criminal use of firearm or electronic defense weapon	5 years
	53a-60c	Assault of an elderly, blind, disabled, pregnant, or mentally retarded person 2nd degree with a firearm	3 years
	53a-60b	Assault or larceny of an elderly, blind, disabled, pregnant, or mentally retarded person 2nd degree	2 years
	53a-217	Criminal possession of firearm or electronic defense weapon	2 years
	53a-60a	Assault 2nd degree with a firearm	1 year
	53a-103a	Burglary 3rd degree with a firearm	1 year
	53a-196f	Possessing child pornography 3rd degree	1 year
	14-223(b)	Increasing speed to elude police after signaling to stop (involving death or serious physical injury)	Subsequent offense: 1 year
	15-154(c)	Refusing to stop boat when ordered by officer in law enforcement vessel and interfering with or endangering a boat, people, or property or increasing speed to escape or elude (causing death or serious physical injury)	Subsequent offense: 1 year

	29-34	Illegal sale or transfer of handgun to person under age 21	1 year
	53-202c	Possession of an assault weapon	1 year
Class A Misdemeanor	53a-61(a)(3)	Assault 3rd degree (when, with criminal negligence, cause physical injury with deadly weapon, dangerous instrument, or electronic defense weapon)	1 year
	53a-61a	Assault of an elderly, blind, disabled, pregnant, or mentally retarded person 3rd degree	1 year
Unclassified	53a-136a	Carjacking	3 years
	53a-300	Acts of terrorism (when commit a class B felony)	10 years (law authorizes the court to impose the penalty for the next most serious degree of felony; if the felony is a class B felony, this law imposes the penalty for a class A felony which carries a 10-year mandatory minimum sentence)
	14-36(h)	Operating a motor vehicle without a license or with a suspended or revoked license	90 days if 2+ prior offenses
	21a-278a(a)	Sale of drugs to minor	2 years (in addition and consecutive to any imprisonment for the underlying drug crime)
	21a-278a(c)	Using person under 18 to sell drugs	3 years (in addition and consecutive to any imprisonment for the underlying drug crime)
	53-202	Commit class A, B, or C felony with assault rifle	8 years (in addition and consecutive to any imprisonment for the underlying drug crime)
	53-202k	Commit class A, B, or C felony with firearm	5 years (in addition and consecutive to any imprisonment for the underlying drug crime)
Classification	**Statute §**	**Offense**	**Presumptive Sentence***

Class A Misdemeanor	21a-267(c)	Use, possession, or delivery of drug paraphernalia near school by non-student	1 year (in addition and consecutive to any imprisonment for the underlying violation)
Unclassified	21a-278(a)	Manufacture or sale of heroin, methadone, cocaine, or crack by non-dependent person	5 years The court may suspend if the person (1) was under age 18 at the time or (2) had significantly impaired mental capacity.
	21a-278(b)	Manufacture or sale of narcotic, hallucinogen, amphetamine, or at least 1 kg marijuana by non-dependent person	1st offense: 5 years Subsequent offense: 10 years The court may suspend if the person (1) was under age 18 at the time or (2) had significantly impaired mental capacity.
	21a-278a(b)	Sale of drugs to minor near school, public housing project, or day care center	3 years (in addition and consecutive to any imprisonment for the underlying drug crime)
	21a-279(d)	Possess narcotic, hallucinogen, or controlled substance near school or day care center	2 years (in addition and consecutive to any imprisonment for the underlying drug crime)
	14-215(c)	Driving during license suspension for DUI or DUI-related offenses	30 days If offense is after 2nd suspension for DUI-related offenses: 120 days If offense is after 3rd or subsequent suspension for DUI-related offenses: 1 year
	14-227a(g)	Operating a motor vehicle under the influence of alcohol or drugs (DWI) (includes snowmobiles and all-terrain vehicles)	1st offense: 48 hours if not given community service 2nd offense: 120 days 3rd and subsequent offenses: 1 year
	15-133	Operating a vessel (boat) under the influence of alcohol or drugs (DWI)	1st offense: 48 hours if not given community service 2nd offense: 120 days 3rd and subsequent offenses: 1 year
	15-156(d)(1)	Operating boat while certificate or right to operate is suspended or revoked for drunken boating or refusing	30 days

		to stop	
	15-156(d)(2)	Operating boat while certificate or right to operate is suspended or revoked for reckless boating 1st or 2nd degree while under the influence	30 days
	29-37(b)	Carry handgun without a permit	1 year

*Judges can depart from presumptive sentence if no one was hurt during the crime and the defendant (1) did not use, attempt, or threaten to use physical force; (2) was unarmed; and (3) did not use, threaten to use, or suggest that he had a deadly weapon or other instrument that could cause death or serious injury. Defendants must show good cause and can invoke this provision only once (CGS § 21a-283a).

Appendix B: Individual Items in the LSI-R Risk Assessment

All items on the LSI-R are measured either in a “yes-no” format or in a “0-3” rating format, based on the following scale.

3: A satisfactory situation with no need for improvement

2: A relatively satisfactory situation, with some room for improvement evident.

1: A relatively unsatisfactory situation with a need for improvement

0: A very unsatisfactory situation with a very clear and strong need for improvement

Items measured in a “0-3” format are denoted with a *. For scoring purposes, ratings of 0 or 1 are converted to a “yes” answer and ratings of 2 or 3 are converted to a “no” answer. “Yes” answers are worth one point each, while “no” answers are worth zero points.

I. Criminal History Component

1. Any prior adult convictions?
2. Two or more prior adult convictions?
3. Three or more prior adult convictions?
4. Three or more present offenses?
5. Arrested under age 16?
6. Ever incarcerated upon conviction?
7. Escape history from a correctional facility?
8. Ever punished for institutional misconduct?
9. Charge laid or probation/parole suspended during prior community service?
10. Official record of assault/violence?

II. Education/Employment Component [reversed coded]

When in labor market:

11. Currently employed?
12. Frequently unemployed?
13. Never employed for a full year?
14. Ever fired?

School or when in school:

15. Less than regular grade 10?
16. Less than regular grade 12?
17. Suspended or expelled at least once?

III. Education/Employment Component [reverse coded] (continued)

Offender completes #18 only if a homemaker or pensioner. Offender completes #18-20 if in school or working. Offender rates 0 for #18-20 if unemployed.

18. Participation/performance?*
19. Peer interactions?*
20. Authority interactions?*

IV. Financial

21. Problems?*
22. Reliance upon social assistance?

V. Family/Marital

23. Dissatisfaction with marital or equivalent situation?*
24. Non-rewarding parental situation?*
25. Non-rewarding situation with other relatives?*
26. Criminal family/spouse?

VI. Accommodation

27. Unsatisfactory?
28. 3 or more address changes last year?*
29. High crime neighborhood?*

VII. Leisure/Recreation

30. Absence of recent participation in an organized activity?
31. Could make better use of time?*

VIII. Companions

32. A social isolate?
33. Some criminal acquaintances?
34. Some criminal friends?
35. Few anti-criminal acquaintances?
36. Few anti-criminal friends?

IX. Alcohol/Drug Problem

37. Alcohol problem, ever?
38. Drug problem, ever?
39. Alcohol problem, currently?*
40. Drug problem, currently?*
41. Law violations?
42. Marital/Family alcohol/drug problem?
43. Interference with school/work?
44. Receiving medical treatment?
45. Other indicators of alcohol/drug problem?

X. Emotional/Personal

46. Moderate interference?
47. Severe interference (active psychosis)?
48. Mental health treatment, past?
49. Mental health treatment, present?
50. Psychological assessment indicated?

XI. Attitudes/Orientations

51. Supportive of crime?*
52. Unfavorable toward convention?*
53. Poor, toward sentence?
54. Poor, toward supervision?

BIBLIOGRAPHY

Abrams, D. S. (2011b). Is pleading really a bargain? *Journal of Empirical Legal Studies, 8*(1), 200-221.

Albonetti, C. A. (1991). An integration of theories to explain judicial discretion. *Social Problems*, *38*(2), 247-266.

Albonetti, C. A. (1997). Sentencing under the federal sentencing guidelines: Effects of defendant characteristics, guilty pleas, and departures on sentence outcomes for drug offenses, 1991-1992. *Law and Society Review*, *31*(4), 789-822.

Albonetti, C. A. (2002). The joint conditioning effect of defendant's gender and ethnicity on length of imprisonment under the federal sentencing guidelines for drug trafficking/manufacturing offenders. *Journal of Gender, Race, and Justice, 6*, 39-73.

Alvarez, A., & Bachman, R. D. (1996). American Indians and sentencing disparity: An Arizona test. *Journal of Criminal Justice*, *24*(6), 549-561.

Anderson, A. L., & Spohn, C. (2010). Lawlessness in the federal sentencing process: A test for uniformity and consistency in sentence outcomes. *Justice Quarterly*, *27*(3), 362-393.

Andrews, D. A. & Bonta, J. (1995). *The level of supervision inventory-revised.* Toronto: Multi-Health Systems.

Andrews, D. A. & Bonta, J. (2000). *The Level of Service Inventory-Revised.* Toronto, Canada: Multi-Health Systems.

Andrews, D. A. & Bonta, J. (2002). *LSI-R The Level of Service Inventory-Revised users' manual.* Toronto: Multi-Health Systems.

Andrews, D. A. & Bonta, J. (2003). *Level of Service Inventory-Revised: U.S. norms manual supplement.* Toronto, Canada: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*(1), 19-52.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency*, *52*(1), 7-27.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica.* Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Arnold, J., & Arnold, L. (2015). Fixing justice in America. Politico Magazine. Available at http://www.politico.com/magazine/story/2015/ 03/criminal-justice-reformcoalition-for-public-safety-116057.html.

Auerhahn, K. (1999). Selective incapacitation and the problem of prediction. *Criminology*, *37,* 703-734.

Ball, J. D. (2006). Is it a prosecutor's world? Determinants of count bargaining decisions. *Journal of Contemporary Criminal Justice, 22*(3), 241-260.

Baumer, E. P. (2013). Reassessing and redirecting research on race and sentencing. *Justice Quarterly*, *30*(2), 231-261.

Beckett, K., & Sasson, T. (2003). *The politics of injustice: Crime and punishment in America*. Sage Publications.

Bedau, H. A. (1964). Death sentences in New Jersey. *Rutgers Law Review, 19,* 1-2.

Belknap, J. (2007). The invisible woman: Gender, crime, and justice (3rd ed.). Belmont, CA: Thompson Wadsworth.

Bergstrom, M., Dermody, F., Chanenson, S. L., & Hyatt, J. (2009). The next era of sentencing reform…revisited. *Federal Sentencing Reporter, 21*(3), 203-210.

Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review, 48*(3), 386-398.

Berk, R. & Bleich, J. (2014). Forecasts of violence to inform sentencing decisions. *Journal of Quantitative Criminology, 30*(1), 79-96.

Berk, R., Sherman, L., Barnes, G., Kurtz, E., & Ahlman, L. (2009). Forecasting murder within a population of probationers and parolees: A high stakes application of statistical learning. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 172*(1), 191-211.

Bernstein, I. N., Kelly, W. R., & Doyle, P. A. (1977). Societal reaction to deviants: The case of criminal defendants. *American Sociological Review*, 743-755.

Bibas, S. (2006). Transparency and participation in criminal procedure. *New York University Law Review*, *81*, 911-966.

Bickle, G. S. & Peterson, R. D. (1991). The impact of gender-based family roles on criminal sentencing. *Social Problems, 38*(3), 372-394.

Bloch, K. R., Engen, R. L., & Parrotta, K. L. (2014). The intersection of race and gender: an examination of sentencing outcomes in North Carolina. *Criminal Justice Studies*, *27*(4), 419-438.

Bonczar, T. P. (2003). *Prevalence of imprisonment in the U.S. population, 1974-2001*. Washington, DC: Bureau of Justice Statistics.

Bonta, J. (1996). Risk-needs assessment and treatment. In Harland, A. T. (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks, CA: Sage Publications.

Bontrager, S., Barrick, K., & Stupi, E. (2013). Gender and sentencing: A meta-analysis of contemporary research. *The Journal of Gender, Race, and Justice, 16*(2), 349-372.

Brennan, P. K., & Spohn, C. (2009). The joint effects of offender race/ethnicity and sex on sentence length decisions in federal courts. *Race and Social Problems*, *1*(4), 200-217.

Britt, C. L. (2000). Social context and racial disparities in punishment decisions. *Justice Quarterly*, *17*(4), 707-732.

Britton, D. M., Jacobsen, S. K., & Howard, G. E. (2017). *The gender of crime*. Rowman & Littlefield.

Burch, T. (2015). Skin color and the criminal justice system: Beyond black-white disparities in sentencing. *Journal of Empirical Legal Studies*, *12*(3), 395-420.

Burgess-Proctor, A. (2006). Intersections of race, class, gender, and crime: Future directions for feminist criminology. *Feminist Criminology*, *1*(1), 27-47.

Bushway, S. D., & Forst, B. (2013). Studying discretion in the processes that generate criminal justice sanctions. *Justice Quarterly*, *30*(2), 199-222.

Bushway, S., Johnson, B. D., & Slocum, L. A. (2007). Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. *Journal of Quantitative Criminology*, *23*(2), 151-178.

Bushway, S. D., & Piehl, A. M. (2001). Judging judicial discretion: Legal factors and racial discrimination in sentencing. *Law and Society Review, 35*, 733-764.

Bushway, S. D., & Piehl, A. M. (2007). Social science research and the legal threat to presumptive sentencing guidelines. *Criminology & Public Policy*, *6*(3), 461-482.

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, *5*(4), 325-370.

Carroll, J. S., & Payne, J. W. (1976). The psychology of the parole decision process: A joint application of attribution theory and information-processing psychology. In Carroll, J. S. & Payne, J. W. (Eds.), *Cognition and Social Behavior (pp 13-32).* Oxford, England: Lawrence Erlbaum.

Carson, E. A. & Anderson, E. (2016). *Prisoners in 2015.* Washington, DC: Bureau of Justice Statistics. US Department of Justice. Retrieved from http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5869.

Casey, P. M., Elek, J. K., Warren, R. K., Cheesman, F., Kleiman, M., & Ostrom, B. (2014). *Offender risk and needs assessment instruments: A primer for courts.* Williamsburg, VA: National Center for State Courts. Retrieved from http://www.ncsc.org/~/media/Microsites/Files/CSI/BJA%20RNA%20Final%20Report_Combined%20Files%208-22-14.ashx

Chambliss, W. J., & Seidman, R. B. (1971). *Law, order, and power*. Reading, MA: Addison-Wesley.

Chammah, M. (2016). *Two parties, two platforms on criminal justice*. The Marshall Project. Retrieved from https://www.themarshallproject.org/2016/07/18/two-parties-two-platforms-on-criminal-justice.

Chenane, J. L., Brennan, P. K., Steiner, B., & Ellison, J. M. (2015). Racial and ethnic differences in the predictive validity of the Level of Service Inventory-Revised among prison inmates. *Criminal Justice and Behavior, 42*(3), 286-303.

Chesney-Lind, M. (1997). *The female offender: Girls, women, and crime.* Thousand Oaks, CA: Sage Publications.

Chiricos, T. G., & Waldo, G. P. (1975). Socioeconomic status and criminal sentencing: An empirical assessment of a conflict proposition. *American Sociological Review*, 753-772.

Clegg, S. & Dunkerley, D. (1980). *Organization, class and control.* Boston: Routledge and Kegan Paul.

Cohen, M. S., Schum, D. A., Freeling, A. N. S., & Chinnis, J. O. (1985). *On the art and science of hedging a conclusion: Alternative theories of uncertainty in intelligence analysis.* Falls Church, VA: Decision Science Consortium.

Connecticut Sentencing Task Force. (2009). *Recommendation for the Connection Sentencing Commission.* Retrieved from https://sentencing.umn.edu/sites/sentencing.umn.edu/files/connecticut_sc_sentencing_task_force_report_2009.pdf

Cooke, D. J. & Michie, C. (2010). Limitations of diagnostic precision and predictive utility in the individual case: A challenge for forensic practice. *Law and Human Behavior*, *34*, 109-175.

Coulson, G., Ilacqua, G., Nutbrown, V., Giulekas, D., & Cudjoe, F. (1996). Predictive utility of the LSI for incarcerated female offenders. *Criminal Justice and Behavior*, *23*(3), 427-439.

Crawford, C., Chiricos, T., & Kleck, G. (1998). Race, racial threat, and sentencing of habitual offenders. *Criminology*, *36*, 481-512.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum, 14*, 538-554.

Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, *43*(6), 1241-1299.

Crew, B. K. (1991). Sex differences in criminal sentencing: Chivalry or patriarchy?. *Justice Quarterly*, *8*(1), 59-83.

Cullen, F. T. & Gendreau, P. (2000). Assessing correctional rehabilitation: Policy, practice, and prospects. *Policies, Processes, and Decisions of the Criminal Justice System, 3*, 109-175.

Cullen, F. T., Jonson, C. L., & Nagin, D. S. (2011). Prisons do not reduce recidivism: The high cost of ignoring science. *The Prison Journal*, *91*(3), 48S-65S.

Curry, T. R., & Corral-Camacho, G. (2008). Sentencing young minority males for drug offenses: Testing for conditional effects between race/ethnicity, gender and age during the US war on drugs. *Punishment & Society*, *10*(3), 253-276.

D'Alessio, S. J., & Stolzenberg, L. (1993). Socioeconomic status and the sentencing of the traditional offender. *Journal of Criminal Justice*, *21*(1), 61-77.

Daly, K. (1989). Neither conflict nor labeling nor paternalism will suffice: Intersections of race, ethnicity, gender, and family in criminal court decisions. *NPPA Journal*, *35*(1), 136-168.

Daly, K., & Bordt, R. L. (1995). Sex effects and sentencing: An analysis of the statistical literature. *Justice Quarterly*, *12*(1), 141-175.

Davis, A. Y. (1981). *Women, race, and class.* New York: Random House.

De Berker, A. O., Rutledge, R. B., Mathys, C., Marshall, L., Cross, G. F., Dolan, R. J., & Bestmann, S. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nature Communications*, *7*, 10996.

Democratic National Committee. (2017). *2016 Democratic party platform.* Retrieved from http://s3.amazonaws.com/uploads.democrats.org/Downloads/2016_DNC_Platform.pdf.

Demuth, S. (2001). The effect of citizenship status on sentencing outcomes in drug cases. *Federal Sentencing Reporter*, *14*(5), 271-275.

Dieterich, W., Mendoza, C., & Brennan, T. (2016). COMPAS risk scales: Demonstrating accuracy equity and predictive parity. Traverse City, MI: *Northpointe Inc*.

Dixon, J. (1995). The organizational context of criminal sentencing. *American Journal of Sociology*, *100*(5), 1157-1198.

Doerner, J. K. & Demuth, S. (2010). The Independent and Joint Effects of Race/Ethnicity, Gender, and Age on Sentencing Outcomes in US Federal Courts. Justice Quarterly, 27(1), 1-27.

Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*, *42*, 449-467.

Drake, E. K., Aos, S., & Miller, M. G. (2009). Evidence-based public policy options to reduce crime and criminal justice costs: Implications in Washington State. *Victims and Offenders*, *4*(2), 170-196.

Dutton, D. B. & Levine, S. (1989). Socioeconomic status and health: Overview, methodological critique, and reformulation. In Bunker, J. P., Gomby, D. S. & Kehrer, B. H. (Eds.), *Pathways to health* (pp. 29-69). Menlo Park, CA: The Henry J. Kaiser Family Foundation.

Dutton, D. G., & Kropp, P.R. (2000). A review of domestic violence risk instruments. *Trauma, Violence, & Abuse, 1*(2), 171-181.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*(5), 383-386.

Eisenstein, J., & Jacob, H. (1977). *Felony justice: An organizational analysis of criminal courts*. Boston: Little, Brown.

Elek, J. K., Warren, R. K., & Casey, P. M. (2015). *Using risk and needs assessment information at sentencing: Observations from ten jurisdictions.* National Center for State Courts' Center for Sentencing Initiatives. Retrieved from http://www.ncsc.org/~/media/Microsites/Files/CSI/RNA%202015/Final%20PEW%20Report%20updated%2010-5-15.ashx

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining delinquency and drug use.* Beverly Hills, CA: Sage.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 75, 643-669.

Englich, B. (2005). Geben Sie ihm doch einfach fünf Jahre!. *Zeitschrift für Sozialpsychologie*, *36*(4), 215-225.

Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin*, *32*(2), 188-200.

Entman, R. M. (1992). Blacks in the news: Television, modern racism and cultural change. *Journalism Quarterly*, *69*(2), 341-361.

Etienne, M. (2009). Legal and practical implications of evidence-based sentencing by judges. *Chapman Journal of Criminal Justice, 1*(1), 43-60.

Fass, T. L., Heilbrun, K., DeMatteo, D., & Fretz, R. (2008). The LSI-R and the COMPAS: Validation data on two risk-needs tools. *Criminal Justice and Behavior*, *35*(9), 1095-1108.

Fischman, J. B., & Schanzenbach, M. M. (2012). Racial disparities under the federal sentencing guidelines: The role of judicial discretion and mandatory minimums. *Journal of Empirical Legal Studies*, *9*(4), 729-764.

Flores, A. W., Bechtel, K., & Lowenkamp, C. T. (2016). False positives, false negatives, and false analyses: A rejoinder to “Machine bias: There’s software used across the country to predict future criminals. And it’s biased against blacks”. *Federal Probation*, *80*, 38.

Flores, A. W., Lowenkamp, C. T., Smith, P., & Latessa, E. J. (2006). Validating the Level of Service Inventory-Revised on a sample of federal probationers. *Federal Probation*, *70*(2), 44-55.

Franklin, C. A., & Fearn, N. E. (2008). Gender, race, and formal court decision-making outcomes: Chivalry/paternalism, conflict theory or gender conflict?. *Journal of Criminal Justice*, *36*(3), 279-290.

Franklin, T. W. (2013). Sentencing Native Americans in US federal courts: An examination of disparity. *Justice Quarterly*, *30*(2), 310-339.

Franklin, T. W. (2017). Sentencing outcomes in US district courts: Can offenders’ educational attainment guard against prevalent criminal stereotypes?. *Crime & Delinquency*, *63*(2), 137-165.

Frazier, C. E. (1979). Appearance, demeanor, and backstage negotiations: Bases of discretion in a first appearance court. *International Journal of Sociology of Law*, *7*(2), 197–209.

Freeman, N. J. (2006). Socioeconomic status and belief in a just world: Sentencing of criminal defendants. *Journal of Applied Social Psychology*, *36*(10), 2379-2394.

Freiburger, T. L. (2010). The effects of gender, family status, and race on sentencing decisions. *Behavioral Sciences and the Law*, *28*(3), 378-395.

French, S. (1986). *Decision theory: An introduction to the mathematics of rationality*. Halsted Press.

Gainey, R. R., Steen, S., & Engen, R. L. (2005). Exercising options: An assessment of the use of alternative sanctions for drug offenders. *Justice Quarterly*, *22*(4), 488-520.

Garland, D. (2002). The risk of risk. In Ericson, R. V. (Ed.) *Risk and morality* (pp 48-86). Toronto: University of Toronto Press.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the “unparalleled” measure of offender risk? A lesson in knowledge cumulation. *Criminal Justice and Behavior*, *29*(4), 397-426.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works!. *Criminology*, *34*(4), 575-608.

Gertner, J. N. (2010). A short history of American sentencing: Too little law, too much law, or just right. *The Journal of Criminal Law and Criminology*, *100*(3), 691-708.

Gertz, M. G., & Price, A. C. (1985). Variables influencing sentencing severity: Intercourt differences in Connecticut. *Journal of Criminal Justice, 13*(2), 131-139.

Gibbs, J. T. (1988). *Young, black, and male in America.* New York: Auburn House.

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association, 49*(268), 732-764.

Gottfredson, D. M. (1961). Comparing and combining subjective and objective parole predictions. *Research Newsletter, 3*, 11-17.

Gottfredson, S. D. & Gottfredson, D. M. (1985). Screening for risk among parolees: Policy, practice, and method. In Farrington, D. P. & Tarling, R. (Eds) *Prediction in criminology*. SUNY Press.

Green, E. (1970). Race, social status, and criminal arrest. *American Sociological Review*, *35*(3), 476-490.

Griffin, T., & Wooldredge, J. (2006). Sex-based disparities in felony dispositions before versus after sentencing reform in Ohio. *Criminology*, *44*, 893-923.

Grusky, D. (Ed.). (2001). Social stratification: Class, race, and gender in sociological perspective. Boulder, CO: Westview Press.

Hagan, J., Nagel, I. H., & Albonetti, C. (1980). The differential sentencing of white-collar offenders in ten federal district courts. *American Sociological Review*, 802-820.

Hannah-Moffat, K. (1999). Moral agent or actuarial subject: Risk and Canadian women's imprisonment. *Theoretical Criminology, 3*(1), 71-94.

Hannah-Moffat, K. (2005). Criminogenic needs and the transformative risk subject: Hybridizations of risk/need in penality. *Punishment & Society, 7*(1), 29-51.

Hannah-Moffat, K. (2009). Gridlock or mutability: Reconsidering "gender" and risk assessment. *Criminology & Public Policy*, *8*(1), 209-219.

Hannah-Moffat, K. (2013). Actuarial sentencing: An "unsettled" proposition. *Justice Quarterly*, *30*(2), 270-296.

Hannah-Moffat, K., & Shaw, M. (2001). *Taking risks: Incorporating gender and culture into the classification and assessment of federally sentenced women in Canada*. Ottawa, Ontario: Status of Women Canada.

Hanson, R. K. (1997). *The development of a brief actuarial risk scale for sexual offense recidivism* (User Report 97-04). Ottawa: Department of the Solicitor General of Canada.

Harcourt, B. E. (2007). *Against prediction*. Chicago: University of Chicago Press.

Harrel, W. A. (1981). The effects of alcohol use and offender remorsefulness on sentencing decisions. *Journal of Applied Social Psychology*, *11*(1), 83-91.

Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments: Evaluating the 'margins of error' of group v. individual predictions of violence. *British Journal of Psychiatry*, *190*, 60-65.

Hartley, R. D., & Tillyer, R. (2012). Defending the homeland: Judicial sentencing practices for federal immigration offenses. *Justice Quarterly*, *29*(1), 76-104.

Hawkins, D. E. (1980). Perceptions of punishment for crime. *Deviant Behavior, 1*, 193-215.

Head, B. W. (2010). Reconsidering evidence-based policy: Key issues and challenges. *Policy and Society*, *29*(2), 77-94.

Hedderman, C. (1990). Effect of defendants' demeanour on sentencing in the Magistrate's Courts. *Research Bulletin*, (29), 32-36.

Hedlund, J. (2015). Bail revalidation: An analysis of weighted release criteria and bail points. Final report. Retrieved from https://www.documentcloud.org/documents/3462040-Revalidation-Final-Report-10-13-15.html

Heilbrun, K. (2009). *Evaluation for risk of violence in adults.* New York: Oxford University Press.

Helms, R., & Jacobs, D. (2002). The political context of sentencing: An analysis of community and individual determinants. *Social Forces*, *81*(2), 577-604.

Hester, R., & Hartman, T. K. (2017). Conditional race disparities in criminal sentencing: A test of the liberation hypothesis from a non-guidelines state. *Journal of Quantitative Criminology*, *33*(1), 77-100.

Hilbe, J. M. (2014). *Modeling count data.* New York: Cambridge University Press.

Hoge, R. D. (2002). Standardized instruments for assessing risk and need in youthful offenders. *Criminal Justice and Behavior, 29*, 380-396.

Holder, E. (2014). *Attorney General Eric Holder speaks at the National Association of Criminal Defense Lawyers 57th Annual Meeting*, Washington, DC: U.S. Department of Justice. Remarks available from https://www.justice.gov/opa/speech/attorney-general-eric-holder-speaks-national-association-criminal-defense-lawyers-57th

Holleran, D., & Spohn, C. (2004). On the use of the total incarceration variable in sentencing research. *Criminology*, *42*(1), 211-240.

Holmes, M. D., Hosch, H. M., Daudistel, H. C., Perez, D. A., & Graves, J. B. (1996). Ethnicity, legal resources, and felony dispositions in two southwestern jurisdictions. *Justice Quarterly*, *13*(1), 11-30.

Holmes, M. D., Smith, B. W., Freng, A. B., & Muñoz, E. A. (2008). Minority threat, crime control, and police resource allocation in the Southwestern United States. *Crime & Delinquency*, *54*(1), 128-152.

Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2006). Exploring the validity of the Level of Service Inventory-Revised with Native American offenders. *Journal of Criminal Justice*, *34*(3), 331-337.

Howard, J. C. (1975). Racial discrimination in sentencing. *Judicature*, *59*, 121-121.

Hyatt, J. M., Chanenson, S. L., & Bergstrom, M. H. (2011). Reform in motion: The promise and perils of incorporating risk assessments and cost-benefit analysis into Pennsylvania sentencing. *Duquesne Law Review*, *49*(4), 707-749.

Irwin, J. (1985). *The jail: Managing the underclass in American society*. Berkeley, CA: University of California Press.

James, D. J., & Glaze, L. E. (2006). *Mental health problems of prison and jail inmates*. Washington, DC: US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

James, N. (2015). *Risk and needs assessment in the criminal justice system.* Congressional Research Service. Retrieved from https://fas.org/sgp/crs/misc/R44087.pdf

Janis, I. L. & Mann, L. (1977). *Decision-making: A psychological analysis of conflict, choice, and commitment.* New York: Free Press.

Johnson, B. D. (2005). Contextual disparities in guidelines departures: Courtroom social contexts, guidelines compliance, and extralegal disparities in criminal sentencing. *Criminology 43*, 761-796.

Johnson, B. D., & Betsinger, S. (2009). Punishing the “model minority”: Asian-American criminal sentencing outcomes in federal district courts. *Criminology*, *47*, 1045-1090.

Johnson, B. D., & DiPietro, S. M. (2012). The power of diversion: Intermediate sanctions and sentencing disparity under presumptive guidelines. *Criminology*, *50*, 811-850.

Johnson, B. D., & King, R. D. (2017). Facial profiling: Race, physical appearance, and punishment. *Criminology*, *55*, 520-547.

Johnson, B.D., King, R. D., & Spohn, C. (2016). Sociolegal approaches to the study of guilty pleas and prosecution. *Annual Review of Law and Social Science*, *12*, 479-495.

Johnson, B. D., & Lee, J. G. (2013). Racial disparity under sentencing guidelines: A survey of recent research and emerging perspectives. *Sociology Compass*, *7*(7), 503-514.

Johnson, E. J., Shu, S. B., Dellaert, B. G., Fox, C., Goldstein, D. G., Häubl, G., ... & Wansink, B. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, *23*(2), 487-504.

Kahneman, D. (2011). *Thinking, fast and slow.* New York, NY: Farrar, Straus, and Girous.

Kalven, H., Jr., & Zeisel, H. (1966). *The American jury*. Boston: Little Brown.

Kauder, N. B., & Ostrom, B. J. (2008). *State sentencing guidelines: Profiles and continuum.* Williamsburg, VA: National Center for State Courts.

Kelly, M. R., Howell, J. C., & Johnson, K. (2013). *Final report for the state of Connecticut: Department of Children and Families, June 2013*. Connecticut Parole Report.

Kennedy, R. (1997). *Race, crime, and the law.* New York: Pantheon Books.

Keren, G. and Gerritsen, L. E. (1999). On the robustness and possible accounts of ambiguity aversion. *Acta Psychologica*, *103*(1), 149-172.

Kern, R. P., & Bergstrom, M. H. (2013). A View from the field: Practitioners' response to actuarial sentencing: An "unsettled" proposition. *Federal Sentencing Reporter*, *25*(3), 185-189.

Keynes, J. M. (1921). *A treatise on probability*. London: Macmillan.

King, N.J., Soule, D.A., Steen, S., & Weidner, R.R. (2005). When process affects punishment: Differences in sentences after guilty plea, bench trial, and jury trial in five guidelines states. *Columbia Law Review, 105,* 959-1009.

King, R. D., & Johnson, B. D. (2016). A punishing look: Skin tone and Afrocentric features in the halls of justice. *American Journal of Sociology*, *122*(1), 90-124.

Kleiman, M., Ostrom, B. J., & Cheesman, F. L. (2007). Using risk assessment to inform sentencing decisions for nonviolent offenders in Virginia. *Crime & Delinquency*, *53*(1), 106-132.

Koons-Witt, B. A. (2002). The effect of gender on the decision to incarcerate before and after the introduction of sentencing guidelines. *Criminology*, *40*(2), 297-328.

Koons-Witt, B. A., Sevigny, E. L., Burrow, J. D., & Hester, R. (2014). Gender and sentencing outcomes in South Carolina: Examining the interactions with race, age, and offense type. *Criminal Justice Policy Review*, *25*(3), 299-324.

Kramer, J., & Steffensmeier, D. (1993). Race and imprisonment decisions. *The Sociological Quarterly*, *34*(2), 357-376.

Kramer, J. H., & Ulmer, J. T. (2009). *Sentencing guidelines: Lessons from Pennsylvania*. Boulder: Lynne Rienner Publishers.

Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*, *28*(4), 471-489.

Kurlychek, M. C., & Johnson, B. D. (2004). The juvenile penalty: A comparison of juvenile and young adult sentencing outcomes in criminal court. *Criminology, 42*, 485-515.

Kutateladze, B. L., Andiloro, N. R., Johnson, B. D., & Spohn, C. C. (2014). Cumulative disadvantage: Examining racial and ethnic disparity in prosecution and sentencing. *Criminology, 52*, 514-551.

Kutateladze, B.L., Lawson, V.Z., & Andiloro, N.R. (2015). Does evidence really matter? An exploratory analysis of the role of evidence in plea bargaining in felony drug cases. *Law and Human Behavior*, *39*(5), 431-442.

La Vigne, N. G., Bieler, S., Cramer, L., Ho, H., Kotonias, C., Mayer, D., ... & Samuels, J. (2014). *Justice reinvestment initiative state assessment report*. Washington, DC: Urban Institute.

LaFrentz, C. D., & Spohn, C. (2006). Who is punished more harshly in federal court? The interaction of race/ethnicity, gender, age, and employment status in the sentencing of drug offenders. *Justice Research and Policy*, *8*(2), 25-56.

Laskorunsky, J. A. (2017). Assessment of offender risk at sentencing: A potential for disparity? In Ulmer, J. T. & Bradley, M. S. (Eds) *Handbook on punishment decisions: Locations of disparity.* Routledge.

Latessa, E. J. & Lovins, B. (2010). The role of offender risk assessment: A policy maker guide. *Victims and Offenders*, *5*(3), 203-219.

Lawrence, A. (2015). *Making sense of sentencing: State systems and policies.* Denver, CO: National Conference of State Legislatures. Retrieved from http://www.ncsl.org/documents/cj/sentencing.pdf.

Lee, J. G. (2016). To detain or not to detain? Using propensity scores to examine the relationship between pretrial detention and conviction. *Criminal Justice Policy Review*. Advance online publication. doi: 10.1177/0887403416668016.

Legislative Program Review & Investigations Committee. (2005). *Mandatory minimum sentences.* Connecticut General Assembly. Retrieved from https://www.cga.ct.gov/2005/pridata/Studies/Mandatory_Minimum_Sentences_Final_Report.htm.

Lippman, W. (1922). *Public opinion.* New York: Harcourt & Brace.

London, C. (2011). Racial impact statements: A proactive approach to addressing racial disparities in prison populations. *Law & Inequality, 29*, 211-248.

Lowenkamp, C. T., Holsinger, A. M., & Latessa, E. J. (2001). Risk/need assessment, offender classification, and the role of childhood abuse. *Criminal Justice and Behavior, 28*, 543-563.

Manchak, S. M., Skeem, J. L., Douglas, K. S., & Siranosian, M. (2009). Does gender moderate the predictive utility of the Level of Service Inventory—Revised (LSI-R) for serious violent offenders?. *Criminal Justice and Behavior*, *36*(5), 425-442.

March, J. G. & Simon, H. . (1958). *Organizations.* Oxford, England: Wiley.

Marcus, M. (2009). The root of the problem: Just deserts and risk assessment. *Florida Law Review, 61*, 751-776.

Mauer, M. (2008). Racial impact statements—Changing policies to address disparities. *Criminal Justice, 23*, 16-21.

Maurutto, P., & Hannah-Moffat, K. (2006). Assembling risk and the restructuring of penal control. *British Journal of Criminology, 46*(3), 438-454.

Mazzella, R., & Feingold, A. (1994). The effects of physical attractiveness, race, socioeconomic status, and gender of defendants and victims on judgments of mock jurors: A meta-analysis. *Journal of Applied Social Psychology*, *24*(15), 1315-1338.

Miethe, T. D., & Moore, C. A. (1985). Socioeconomic disparities under determinate sentencing systems: A comparison of preguideline and postguideline practices in Minnesota. *Criminology*, *23*(2), 337-363.

Mihailides, S., Judge, B., & Van den Bossche, E. (2005). The LSI-R in an Australian setting: Implications for risk/needs decision-making in forensic contexts. *Psychiatry, Psychology, & Law, 12*, 207-217.

Miller, E. (1986). *Street women.* Philadelphia, PA: Temple University Press.

Millson, B., Robinson, D., Rubin, M., & Van Dieten, M. (2009). *A process evaluation of the women offender case management model, implemented by the Connecticut Court Support Services Division.* Orbis Partners, Inc. Retrieved from http://www.ct.gov/opm/lib/opm/cjppd/cjcjpac/process_report_women_offenders.pdf

Mitchell, O. (2005). A meta-analysis of race and sentencing research: Explaining the inconsistencies. *Journal of Quantitative Criminology*, *21*(4), 439-466.

Monahan, J. & Skeem, J. L. (2014). Risk redux: The resurgence of risk assessment in criminal sanctioning. *Federal Sentencing Reporter, 26*, 158-166.

Monahan, J. & Skeem, J. L. (2015). Risk assessment in criminal sentencing. *Annual Review of Clinical Psychology, 12*, 489-513.

Monahan, J., Skeem, J., & Lowenkamp, C. (2017). Age, risk assessment, and sanctioning: Overestimating the old, underestimating the young. *Law and Human Behavior*, *41*(2), 191-201.

Monat, A., Averill, J. R., & Lazarus, R. S. (1972). Anticipatory stress and coping reactions under various conditions of uncertainty. *Journal of Personality and Social Psychology*, *24*(2), 237-253.

Mustard, D. B. (2001). Racial, ethnic, and gender disparities in sentencing: Evidence from the US federal courts. *The Journal of Law and Economics*, *44*(1), 285-314.

Myers, M. A. (1987). Economic inequality and discrimination in sentencing. *Social Forces*, *65*(3), 746-766.

Nagel, I. H., & Hagan, J. (1983). Gender and crime: Offense patterns and criminal court sanctions. *Crime and Justice*, *4*, 91-144.

Nagel, S. (1969). *The legal process from a behavioral perspective*. Homewood, IL: The Dorsey Press.

Nobiling, T., Spohn, C., & DeLone, M. (1998). A tale of two counties: Unemployment and sentence severity. *Justice Quarterly*, *15*(3), 459-485.

Nowacki, J. S. (2017). An intersectional approach to race/ethnicity, sex, and age disparity in federal sentencing outcomes: An examination of policy across time periods. *Criminology & Criminal Justice*, *17*(1), 97-116.

Oleson, J. C. (2011). Risk in sentencing: Constitutionally suspect variables and evidence-based sentencing. *Southern Methodist University Law Review*, *64*, 1329-1402.

Oliver, M. B. (2003). African American men as "criminal and dangerous": Implications of media portrayals of crime on the "criminalization" of African American men. *Journal of African American Studies*, *7*(2), 3-18.

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, *26*, 156-176.

Ostrom, B., Kleiman, M., Cheesman, F., Hansen, A. M., & Kauder, N. B. (2002). *Offender risk assessment in Virginia: A three-stage evaluation*. Washington, DC: US Department of Justice.

Peters, T. W. & Warren, R. K. (2006). *Getting smarter about sentencing: NCSC's Sentencing Reform Survey*. Williamsburg, VA: National Center for State Courts.

Pew Center on the States. (2011). *Risk/needs assessment 101: Science reveals new tools to manage offenders*. Public Safety Performance Project Issue Brief. Retrieved from http://www.pewtrusts.org/en/researchand-analysis/issue-briefs/2011/09/20/riskneeds-assessment-101-science-reveals-new-tools-to-manageoffenders

Pew-MacArthur Results First Initiative. 2014. *Evidence-based policymaking. A guide for effective government*. Washington, Chicago: The Pew Charitable Trusts & John D. and Catherine T. MacArthur Foundation. Retrieved from http://www.pewtrusts.org/~/media/assets/2014/11/evidencebasedpolicymakingaguideforeffectivegovernment.pdf

Pratt, T. C. (1998). Race and sentencing: A meta-analysis of conflicting empirical research results. *Journal of Criminal Justice*, *26*(6), 513-523.

Quinney, R. (1973). *Critique of the legal order: Crime control in a capitalist society.* Boston: Little Brown.

Ratansi, S., & Cox, S. M. (2007). Assessment and validation of Connecticut's Salient Factor Score. Central Connecticut State University, Connecticut Statistical Analysis Center. Retrieved from http://www.ct.gov/doc/lib/doc/pdf/RevalidationStudy2007.pdf

Rehavi, M. M., & Starr, S. B. (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy*, *122*, 1320-1354.

Reinhart, C. (2008). *Crimes with mandatory minimum sentences—updated and revised*. Office of Legislative Research Report 2008-R-0619. Hartford, CT. Retrieved from https://www.cga.ct.gov/2008/rpt/2008-R-0619.htm

Richey-Mann, C., Zatz, M., & Rodriguez, N. (2006). *Images of color, images of crime: Readings*. Los Angeles: Roxbury.

Robinson, D. T., Smith-Lovin, L., & Tsoudis, O. (1994). Heinous crime or unfortunate accident? The effects of remorse on responses to mock criminal confessions. *Social Forces*, *73*(1), 175-190.

Rodriguez, S. F., Curry, T. R., & Lee, G. (2006). Gender differences in criminal sentencing: Do effects vary across violent, property, and drug offenses?. *Social Science Quarterly*, *87*(2), 318-339.

Roehl, J., & Guertin, K. (2000). Intimate partner violence: The current use of risk assessments in sentencing offenders. *Justice System Journal*, *21*(2), 171-198.

Ruback, R. B., Kempinen, C. A., Tinik, L. A., & Knoth, L. K. (2016). Communicating risk information at criminal sentencing in Pennsylvania: An experimental analysis. *Federal Probation*, *80*(2), 47-56.

Ruth, H. S., & Reitz, K. R. (2003). *The challenge of crime: Rethinking our response*. Cambridge, MA: Harvard University Press.

Sacks, H. R. (1976). Promises, performance, and principles: An empirical study of parole decisionmaking in Connecticut. *Connecticut Law Review*, *9*(3), 347-423.

Sacks, M., & Ackerman, A. R. (2014). Bail and sentencing: Does pretrial detention lead to harsher punishment?. *Criminal Justice Policy Review*, *25*(1), 59-77.

Sampson, R. J. & Wilson, W. J. (1995). Toward a theory of race, crime, and urban inequality. In Hagan, J. & Peterson, R. (Eds.), *Crime and inequality*. Stanford, CA: Stanford University Press.

Santos Silva, J. M. C., and Tenreyo, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641-658.

Shaver, K. G. (1975). *An introduction to attribution process.* Cambridge, MA: Winthrop.

Shermer, L. O. N., & Johnson, B. D. (2010). Criminal prosecutions: Examining prosecutorial discretion and charge reductions in US federal district courts. *Justice Quarterly*, *27*(3), 394-430.

Silver, E., & Miller, L. L. (2002). A cautionary note on the use of actuarial risk assessment tools for social control. *Crime & Delinquency*, *48*(1), 138-161.

Simon, H. A. (1957). *Models of man.* New York: Wiley.

Skeem, J. L. & Lowenkamp, C. T. (2016). Risk, race, and recidivism: predictive bias and disparate impact. *Criminology*, *54*, 680-712.

Skeem, J. L. & Monahan, J. (2011). Current directions in violence risk assessment. *Current Directions in Psychological Science*, *20*(1), 38-42.

Skeem, J. L., Monahan, J., & Lowenkamp, C. (2016). Gender, risk assessment, and sanctioning: The cost of treating women like men. *Law and Human Behavior*, *40*(5), 580-593.

Skolnick, J. (1966). *Justice without trial.* New York: John Wiley.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, *28*(1), 1-39.

Smith, P., Cullen, F. T., & Latessa, E. J. (2009). Can 14,737 women be wrong? A meta-analysis of the LSI-R and recidivism for female offenders. *Criminology & Public Policy*, *8*(1), 183-208.

Sommeiller, E., Price, M., & Wazeter, E. (2016). *Income inequality in the U.S. by state, metropolitan area, and county.* Washington, DC: Economic Policy Institute. Retrieved from http://www.epi.org/files/pdf/107100.pdf.

Spohn, C. (2000). Thirty years of sentencing reform: The quest for a racially neutral sentencing process. *Criminal Justice*, *3*, 427-501.

Spohn, C. (2008). Race, sex, and pretrial detention in federal court: Indirect effects and cumulative disadvantage. *University of Kansas Law Review.*, *57*, 879-901.

Spohn, C., & Beichner, D. (2000). Is preferential treatment of female offenders a thing of the past? A multisite study of gender, race, and imprisonment. *Criminal Justice Policy Review*, *11*(2), 149-184.

Spohn, C. and DeLone, M. (2000). When does race matter?: An analysis of the conditions under which race affects sentence severity. *Sociology of Crime Law and Deviance*, *2*, 3-37.

Spohn, C., Gruhl, J., & Welch, S. (1981). The effect of race on sentencing: A re-examination of an unsettled question. *Law and Society Review*, 71-88.

Spohn, C., & Holleran, D. (2000). The imprisonment penalty paid by young, unemployed black and Hispanic male offenders. *Criminology*, *38*, 281-306.

Starr, S. B. (2014a). Estimating gender disparities in federal criminal cases. *American Law and Economics Review*, *17*(1), 127-159.

Starr, S. B. (2014b). Evidence-based sentencing and the scientific rationalization of discrimination. *Stanford Law Review, 66*, 803-872.

Starr, S. B. (2014c). *Sentencing by the numbers.* New York Times. Retrieved from https://www.nytimes.com/2014/08/11/opinion/sentencing-by-the-numbers.html

State of Connecticut Judicial Branch. (2017). *Connecticut's courts.* Retrieved from https://www.jud.ct.gov/Publications/es201.pdf.

Steffensmeier, D. & Demuth, S. (2001). Ethnicity and judges' sentencing decisions: Hispanic-black-white comparisons. *Criminology, 39*, 145-178.

Steffensmeier, D., Kramer, J. & Streifel, C. (19f93). Gender and imprisonment decisions. *Criminology, 31*(3), 411-446.

Steffensmeier, D., Kramer, J., & Ulmer, J. (1995). Age differences in sentencing. *Justice Quarterly*, *12*(3), 583-602.

Steffensmeier, D., Painter-Davis, N., & Ulmer, J. (2016). Intersectionality of race, ethnicity, gender, and age on criminal punishment. *Sociological Perspectives, 60*(4), 810-833.

Steffensmeier, D., Ulmer, J., & Kramer, J. (1998). The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black, and male. *Criminology, 36*, 763-798.

Stuart, J. & Sykora, R. (2010). Minnesota's failed experience with sentencing guidelines and the future of evidence-based sentencing. *William Mitchell Law Review, 37*, 426-468.

Sumarwan, U., & Hira, T. K. (1993). The effects of perceived locus of control and perceived income adequacy on satisfaction with financial status of rural households. *Journal of Family and Economic Issues*, *14*(4), 343-364.

Thaler, R. & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* New Haven, CT: Yale University Press.

Thornberry, T. P., Lizotte, A. J., Krohn, M. D., Farnworth, M., & Jang, S. J. (1994). Delinquent peers, beliefs, and delinquent behavior: A longitudinal test of interactional theory. *Criminology*, *32*(1), 47-83.

Thornberry, T. P., Moore, M., & Christenson, R. L. (1985). The effect of dropping out of high school on subsequent criminal behavior. *Criminology*, *23*, 3-18.

Tonry, M. (1995). *Malign neglect: Race, crime, and punishment in America*. Oxford University Press.

Tonry, M. (1996). *Sentencing matters*. New York: Oxford University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science,* 185(4157), 1124-1131.

Tyler, T. R. (1990). *Why people obey the law: Procedural justice, legitimacy, and compliance.* New Haven, CT: Yale University Press.

Tyler, T. R. (2003). Procedural justice, legitimacy, and the effective rule of law. *Crime and Justice*, *30*, 283-357.

U.S. Census Bureau. (2010). *Quickfacts: Connecticut.* Retrieved from https://www.census.gov/quickfacts/CT.

U.S. Census Bureau. (2012). Table 16: Resident population by age and state: 2010. *Statistical Abstract of the United States* (131st edition). Retrieved from https://www2.census.gov/library/publications/2011/compendia/statab/131ed/tables/pop.p

U.S. Census Bureau. (2017). Table H-8B: Median income of households by state using three-year moving averages: 1984-2016. *Historical Income Tables: Households.* Retrieved from https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html.

U.S. Department of Health and Human Services. (1999). *Mental health: Culture, race, and ethnicity. A supplement to Mental health: A report of the Surgeon General.* Washington, DC: U.S. Public Health Services.

U.S. Department of Justice, Federal Bureau of Investigation. (2009). Crime in the United States, 2009: Table 35 - Five-Year Arrest Trends, by Sex 2005-2009. Retrieved September 28, 2011, from http://www2.fbi.gov/ucr/cius2009/data/table_35.html

Ulmer, J. T. (2012). Recent developments and new directions in sentencing research. *Justice Quarterly*, *29*(1), 1-40.

Ulmer, J. T., & Johnson, B. (2004). Sentencing in context: A multilevel analysis. *Criminology*, *42*(1), 137-178.

Ulmer, J. T., Kurlychek, M. C., & Kramer, J. H. (2007). Prosecutorial discretion and the imposition of mandatory minimum sentences. *Journal of Research in Crime and Delinquency*, *44*(4), 427-458.

Ulmer, J., Painter-Davis, N., & Tinik, L. (2016). Disproportional imprisonment of Black and Hispanic males: Sentencing discretion, processing outcomes, and policy structures. *Justice Quarterly*, *33*(4), 642-681.

Van Nostrand, M. & Lowenkamp, C. T. (2013). *Assessing pretrial risk without a defendant interview.* Houston, TX: Laura and John Arnold Foundation. Retrieved from http://www.arnoldfoundation.org/sites/default/files/pdf/LJAF_Report_no-interview_FNL.pdf

Von Hirsch, A. (1976). *Doing justice: The choice of punishments*. New York: Hill & Wang.

Warren, R. K. (2007). Evidence-based practices and state sentencing policy: Ten policy initiatives to reduce recidivism. *Indiana Law Journal, 82*(5), 1307-1318.

Warren, R. K. (2008). Evidence-based sentencing: The application of principles of evidence-based practice to state sentencing practice and policy. *University of San Francisco Law Review*, *43*, 585-634.

Welch, K., Payne, A. A., Chiricos, T., & Gertz, M. (2011). The typification of Hispanics as criminals and support for punitive crime control policies. *Social Science Research*, *40*(3), 822-840.

Western, B. (2006). *Punishment and inequality in America.* Russell Sage Foundation.

Wheeler, S., Weisburd, D., & Bode, N. (1982). Sentencing the white-collar offender: Rhetoric and reality. *American Sociological Review*, 641-659.

Whiteacre, K. W. (2006). Testing the Level of Service Inventory–Revised (LSI-R) for racial/ethnic bias. *Criminal Justice Policy Review 17*(3), 330-342.

Willick, D. H., Gehlker, G., & McFarland Watts, A. (1975). Social class as a factor affecting judicial disposition: Defendants charged with criminal homosexual acts, *Criminology*, *13*(1), 57-77.

Wolfe, S. E., Pyrooz, D. C., & Spohn, C. C. (2011). Unraveling the effect of offender citizenship status on federal sentencing outcomes. *Social Science Research*, *40*(1), 349-362.

Wolff, M.A. (2008). Evidence-based judicial discretion: Promoting public safety through state sentencing reform. *New York University Law Review, 83*(5), 1389-1419.

Wong, S. C., & Gordon, A. (2006). The validity and reliability of the Violence Risk Scale: A treatment-friendly violence risk assessment tool. *Psychology, Public Policy, and Law*, *12*(3), 279.

Wooldredge, J. (2010). Judges' unequal contributions to extralegal disparities in imprisonment. *Criminology, 48*, 539-567.

Wooldredge, J., Frank, J., Goulette, N., & Travis, L. (2015). Is the impact of cumulative disadvantage on sentencing greater for black defendants?. *Criminology & Public Policy*, *14*(2), 187-223.

Wu, J., & Spohn, C. (2009). Does an offender's age have an effect on sentence length? A meta-analytic review. *Criminal Justice Policy Review*, *20*(4), 379-413.

Yang, M., Wong, S.C.P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*(5), 740-767.

Zatz, M. S. (1987). The changing forms of racial/ethnic biases in sentencing. *Journal of Research in Crime and Delinquency*, *24*(1), 69-92.

Zatz, M. S. (2000). The convergence of race, ethnicity, gender, and class on court decision making: Looking toward the 21st century. *Criminal Justice*, *3*(1), 503-552.

Zeisel, H., & Diamond, S. S. (1977). Search for sentencing equity: Sentence review in Massachusetts and Connecticut. *Law & Social Inquiry*, *2*(4), 881-932.

CASE LAW CITED

Malenchik v. State, (2010), 908 N.E.2d 710 (Ind.)

State v. Loomis, (2016), 881 N.W.2d 749 (Wisc.)

Williams v. New York, (1949), 337 U.S. 241

Woodson v. North Carolina, (1976), 428 U.S. 280