

Gender Disparity and Ecological Contexts of Court Community
in Federal District Courts

by

Byung Bae Kim

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2015 by the
Graduate Supervisory Committee;

Cassia Spohn, Chair
Xia Wang
Kevin Wright

ARIZONA STATE UNIVERSITY

August 2015

ABSTRACT

Gender disparity in sentencing outcomes has a long tradition in sentencing literature, with a substantial body of evidence indicating that women offenders are treated with greater leniency over male counterparts. The prior literature on gender and sentencing, however, has ignored broader social contexts within which judicial decision-making occurs. This dissertation attempts to address this limitation by dissecting the nature of gender disparity through ecological lenses. Using federal sentencing data for fiscal year 2001 through 2010 and other complementary data sets, this dissertation, divided into two major sub-studies, has examined the roles of two social contextual variables, such as religious and political conservatism, in producing gender differentials in sentencing outcomes.

The key findings revealed in this dissertation are as follows. The first study has demonstrated that 1) the impact of religious and political conservatism has indeed reduced the level of gender disparity, with the female discount in sentencing outcomes dissipating in court communities characterized by higher levels of religious and political conservatism and 2) the conditioning effects of both religious and political conservatism has turned out to have racial implications, as African American female offenders were more likely to be influenced by religious and political conservatism, relative to their white counterparts. As an extension of the first study, the second study has sought to assess the impact of policy changes related to gender through the ecological contexts of court communities. The results of this second study have shown that 3) the impact of federal guidelines' transition into advisory guidelines, especially the *Gall v. U.S.* decision,

has increased gender disparity in sentencing outcomes in federal district courts and 2) the increase has been greater in court communities represented by lower levels of religious and political conservatism. Overall, this dissertation has added to literature by demonstrating the possibility that gender disparity is deeply entrenched in the ecological contexts of court communities. Finally, I have discussed the main implications of the findings, as well as future directions.

ACKNOWLEDGEMENTS

This project would not have been possible without help and guidance from the following persons. First and foremost, my family has provided a continuous support for my studies at ASU. My wife, Jungmi, has sacrificed so much of her career to help me finish this project. My two daughters, Yujung and Minjung, had to endure the absence of their dad for many years. I would also be remiss in not extending my gratitude to my parents, sister and brother for their endless support and love.

Over the past 5 years, I have learned a great deal from the world's renowned faculty members at ASU. Most importantly, I would like to extend my sincere appreciation to my chair, Dr. Cassia Spohn. Her expertise and patience greatly helped me finish this project and she has always been a great source of guidance since the first day of my graduate education. I also want to thank my two other committee members, Dr. Xia Wang and Dr. Kevin Wright, who have always welcomed me and never minded my dumb questions. Finally, I also thank my mentors from South Korea, Dr. Baikchul Lee, Dr. Okkyung Yoon, Dr. Heejong Joo, Dr. Suhong Min, and Dr. Soojung Lee for their constant encouragement and inspiration at various stages of my life.

Last, but certainly not least, I want to say thanks to my dear friends in the school of Criminology and Criminal Justice, who provided warm friendships, many laughs, and sometimes serious discussions over the years. I want to thank Mario Cano, Janne Gaub, Lisa Dario, Weston Morrow, Marcus Galaste, Richard Moule, Jr, Samuel Vickovic, and Rachel Cheon for sharing important parts of their lives with me. My special thanks goes out to KiDeuk Kim who has provided sincere friendship and professional insights.

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
LIST OF FIGURES	iv
CHAPTER	
1 INTRODUCTION	1
2 LITERATURE REVIEW	9
Sentencing Reforms and the Federal Experiment	9
The History of the 20 th Century Sentencing Reforms.....	9
The Case of Federal Reforms: Focusing on the U.S. Sentencing Guidelines ...	14
Gender Disparity in Sentencing.....	26
Historical Contexts	26
Current Evidence on Gender Equality in Sentencing	29
Theoretical Approaches to Gender Disparity in Sentencing	34
The Importance of Gender Norms in Sex-Based Disparity	40
The Impact of Ecological Contexts in Sentencing	46
Theoretical Frameworks: Focusing on Court Community Perspective	47
Empirical Evidence of Key Contextual Variables	54
The Case of Jurisdiction Variation in Federal District Courts.....	59
Summary of Prior Literature	70
3 CURRENT FOCUS	74
STUDY-1: Gender Disparity and The Impact of Ecological Contexts:	
Implications on Racial and Ethnic Heterogeneity	74

CHAPTER	Page
Religious Context and Gender Disparity	75
Political Context and Gender Disparity	78
Gender Disparity and Ecological Contexts along Racial/Ethnic Lines	82
Hypotheses	84
STUDY-2: Gender Disparity, Policy Changes, and Ecological Contexts of Court Community	87
Gender Disparity and Changes in Sentencing Policies	88
Gender Disparity, Booker and Social Contexts	91
Hypotheses	95
4 METHODOLOGY	99
Data	99
Dependent Variable	100
Key Independent Variables	100
Control Variables	101
Common Analytic Issues	106
Two-stage v. One-stage model	106
Departure and mandatory minimum status as controls	109
Missing Data	111
STUDY-1	112
Data	112
Measures	112
Analytic Strategy.....	113

CHAPTER	Page
STUDY-2	115
Data	115
Measures	115
Analytic Strategy.....	118
5 ANALYSES AND RESULTS	121
STUDY-1	121
Main Hypotheses	121
Descriptive Statistics	122
Bivariate Relationships	125
Unconditional Model	129
Sentence Length Decision: Hypothesis-1 through -4	130
Supplemental Analyses	135
Race/Ethnicity-Specific Results: Hypothesis-5 and -6	137
Summary of the Main Findings	141
STUDY-2	142
Main Hypotheses	142
Descriptive Statistics	143
Unconditional Model	145
Results regarding Hypothesis-1 and -2	147
Supplemental Analyses	150

CHAPTER	Page
Results regarding Hypotheses-3 and -4	152
Summary of the Main Findings	158
6 DISCUSSION AND CONCLUSION	159
Discussion of the Main Findings	159
STUDY-1	159
STUDY-2	165
Strengths and Limitations of the Present Study	169
Theoretical and Policy Implications	173
Conclusion	181
REFERENCES	183

LIST OF TABLES

Table	Page
1 Variables Description: Codes and Summary	105
2 Descriptive statistics	124
3 Correlation Matrix for Individual-Level Variables	126
4 Correlation Matrix for District-Level Variables	127
5 Unconditional Model	130
6 Sentence Length Models	132
7 Model Comparisons by Departure Controls	136
8 Cross-level Interaction between Gender and Religion by Race/Ethnicity	138
9 Cross-level Interaction between Gender and Politics by Race/Ethnicity	140
10 Descriptive statistics	144
11 Unconditional Model	146
12 Full Models	148
13 Model Comparisons by Mandatory Minimum Controls	151
14 Three-way Cross-level Interactions	153

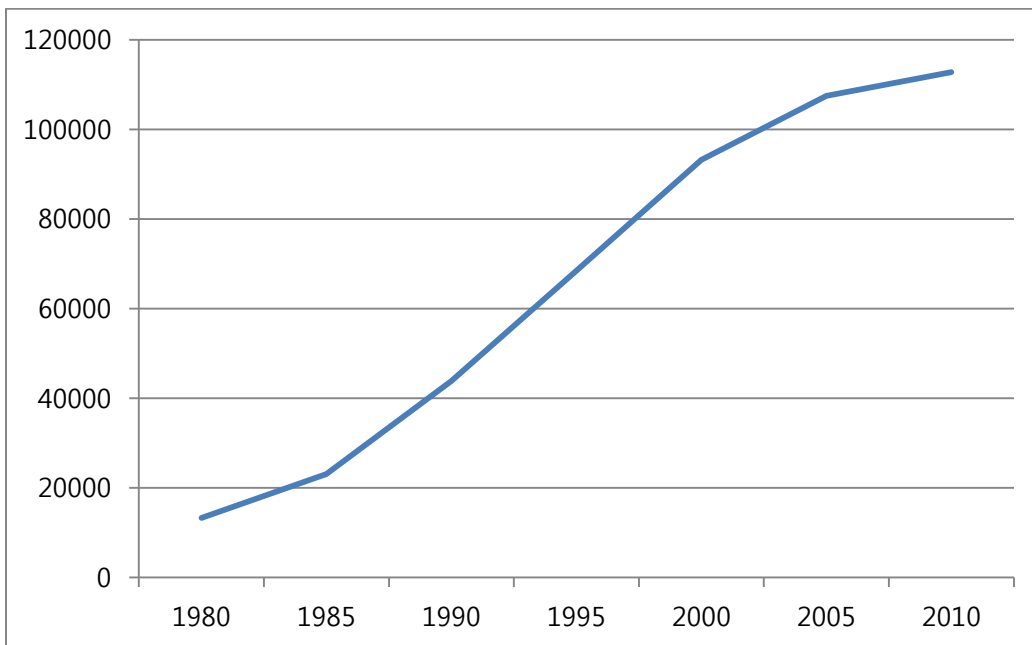
LIST OF FIGURES

Figure	Page
1 Female Incarceration Rates (1980-2010)	1
2 Federal Sentencing Table	18
3 Federal Court Structures	60
4 Conceptual Framework related to Hypotheses-1 and -2	96
5 Booker and Gall Variables Coding	117
6 Scatterplot for Religious and Political Conservatism	129
7 Cross-level interactions between gender and religious/political conservatism ..	134
8 Cross-level interaction between gender and religious conservatism by race	139
9 Cross-level interaction between gender and political conservatism by race	141
10 Interaction between gender and Booker/Gall	150
11 Three-way interaction-religious conservatism, gender, Booker/Gall	155
12 Three-way interaction-political conservatism, gender, Booker/Gall	157
13 Multiple layers surrounding federal judges	178

CHAPTER 1: INTRODUCTION

Traditionally, women in the criminal justice system were considered the forgotten few, as they received little attention due to the lower risk they posed to society and their small numbers (Pollack, 1950). In addition, it has been generally known that female offenders are treated with greater leniency over their male counterparts (Daly & Bordt, 1995), even though there were some historical exceptions to this general social phenomenon (see Boritch, 1993). However, there has been a dramatic growth in women's imprisonment over the past few decades (as shown in figure-1) and it has been estimated that about one million women are currently under some form of criminal justice system control (Mauer, 2013).

Figure-1 Female Incarceration Rates (1980-2010)



Source: Carson & Golinelli (2013)

Critics argue that both an increase in offending rates among women and changes in criminal justice policies, such as different law enforcement strategies and sentencing policies, are largely responsible for these unexpected changes directed toward female offenders (Chesney-Lind, 1991; Kruttschnitt & Gartner, 2003; Mauer, Potler, & Wolf, 1999). Among others, the abrupt shift in sentencing policies—what is better known as the sentencing reform movement in the 1970s and 1980s—is the one cited most often in the literature for the expansion of women under the control of the criminal justice system (Chesney-Lind, 1991).

As the main determinant of the get-tough tendency on women offenders, the primary goals of the sentencing reforms over the past few decades are to reduce extra-legal disparity and increase uniformity in sentencing, thus providing more equal treatment in courtrooms (Frankel, 1972; Stith & Cabranes, 1998). The architects of the sentencing reform movement attempted to attain these goals by constraining judicial discretion through structured sentencing schemes, such as sentencing guidelines, mandatory minimums, and determinate sentencing (Tonry, 1996). Reducing racial and ethnic disparity in sentencing outcomes was one of the most salient driving forces behind sentencing reform, as there was a near consensus on this issue (Spohn, 2000). Gender equality, however, has garnered distinctively different attention because the concept of gender disparity in sentencing has posed a “somewhat unique problem” (Blumstein, Cohen, Martin, & Tonry, 1983, p. 114). This is because, unlike racial disparity in sentencing, in which the less powerful groups in society are disadvantaged in courtrooms relative to their counterparts, women, who are considered less powerful and own less

economic and political power, had been treated more leniently in criminal courts, as compared to men (Curran, 1983). Put differently, given the judicial leniency toward female defendants in the pre-reform era, reducing sex-based disparities generally suggests a substantial increase in punishment handed down to women (Nagel & Johnson, 1994). In this vein, critics claim that the recent law-and-order movement, represented by the war on drugs and the mass incarceration binge, actually disadvantaged more women than men (e.g., Chesney-Lind, 1995). Similarly Daly and Tonry (1997) also contended that women in the post-reform era face more severe but equal punishment to their male counterparts “in the name of a restricted notion of equality with men” (p. 241).

Against this backdrop, a group of scholars have paid renewed attention to the issue of whether and under what conditions female defendants are treated more leniently at sentencing (Spohn, 1999). Overall, prior studies have revealed that, when compared to their male counterparts, females generally experience preferential treatment even in the post-reform era (See Bontrager, Barrick, & Stupi, 2013; Daly & Bordt, 1995), and, to a lesser degree, the effect of gender is also conditioned by individual circumstances, such as types of crime (Rodriguez, Curry & Lee, 2006), criminal history (Spohn, 1998), and familial circumstances (Daly, 1987a, 1989a) of female defendants. Finally, a small body of studies also suggests that gender disparity has not dramatically changed across the sentencing reform movement (Griffin & Wooldredge, 2006; Koons-Witt, 2002).

Although this body of research has significantly advanced scholarship on gender-based sentencing disparity, it also needs to be noted that the extant work has limited its

attention to the immediate circumstances surrounding female defendants, thus neglecting the broader socio-political contexts within which judicial decisions are made. This oversight is quite *odd* given the well-established tradition in the field of sentencing in which racial disparity is analyzed through the contextual lenses (See, Britt, 2000; Kautt, 2002; Myers & Talarico, 1987; Ulmer & Johnson, 2004; Wang & Mears, 2010a, 2010b). Indeed, the lack of interest in contextual factors in conjunction with gender is regrettable, and studies focusing on the influences of the ecological context of the court community are sorely needed for at least two reasons. First, such research may help to reconcile discrepancies in the previous findings on gender disparity and thus contribute to generalizability of the findings related to gender and gender-related factors on sentencing (Daly & Bordt, 1995; Helms & Jacobs, 2002). Second, the attempt to understand gender disparity in terms of ecological contexts would allow us to dissect the nature of gender disparity, thus providing important insights on *why* females are at an advantage over males (Nagel & Hagan, 1983).

Importantly, the focal concerns perspective, a dominant theoretical approach to accounting for sentencing disparity, identifies societal stereotypes on gender as the key determinant driving gender disparity in courtroom decision-making (Steffensmeier, Ulmer, & Kramer, 1998). If that argument is correct, then one should observe a systematic gender disparity depending on the level of gender-based role expectation across different court communities (See Daly & Bordt, 1995; Kruttschnitt & Green, 1984). Indeed, the possibility that gender effects may vary across court communities could be gleaned from Pollack (1950). More than six decades ago, Pollack noted that

chivalry—the notion that female offenders are treated more leniently than male offenders—is derived from traditional role-playing shaped by cultural patterns that reflect status differences between the sexes (Pollack, 1950). It can be further inferred that both the cultural patterns and status differences could vary across ecological contexts, which in turn influence substantive rationality formed within court communities. Taken as a whole, then, although this line of research could significantly advance our understanding on gender disparity, little systemic effort has been made to explore if gender disparity is contextualized.

Reflecting this concern, in the first part of this dissertation, I will assess the contingent effects of two interrelated, but distinct social contexts—local religious and political environments— and examine whether these social contexts moderate the impact of gender on sentencing outcomes. Here I specifically focus on the religious and political contexts, because they are suggested to be closely related to both gender stereotypes and punishment severity in prior studies (Baumer & Martin, 2013; King, 2008; Ulmer, Bader, & Gault, 2008). Indeed, one of the leading authorities in the U.S. on sentencing policies, Michael Tonry, underscored the importance of the two social contexts, claiming that conservative politics and fundamentalist religious views were the two most critical social forces shaping the U.S. sentencing policy and practice (Tonry, 2009). Therefore, in this part of the dissertation, I will examine the question of whether an offender’s gender may be especially influential in some courts characterized by distinct political and religious conservatisms, such that political and religious conservatism may amplify gender-based sentencing disparity (Helms & Jacobs, 2002; Kruttschnitt & Green, 1984). While

exploring the relationship between gender and conservatism, I will also investigate the possibility that the potential moderating effects of the ecological contexts are further colored by the race and ethnicity of defendants. As claimed elsewhere, the relationship among community contexts, crime occurrence, and the race/ethnicity of offenders is argued to be inseparable (Sampson, 2012; Wilson, 1987). At the same time, it has also been claimed that there exist distinct gender stereotypes for white, Black, and Hispanic women (Brennan, 2006; Krivo, Paterson, & Hagan, 2006). Combining these two lines of research, I will examine whether chivalry is extended to minority females only in some court environments, thus shedding renewed light on the long discussed issue of whether chivalry bypasses women of color (Belknap, 2001).

As an extension of the first study, in the second part of the dissertation I will expand on the main argument set forth earlier—that is, gender disparity is shaped by the ecological contexts of court communities. The main thesis I will advance in the latter half of the dissertation is that the degree to which gender disparity is impacted by formal policies attempting to modify the disparity may not be invariant across court communities. Instead, the potential changes may differ across the ecological contexts of court communities. The court community perspective, a dominant theoretical framework for social ecology and sentencing, states that the effect of policy changes is almost always filtered through so-called *informal sanction norms*, which are essentially shaped by the social ecology of the court community (Eisenstein, Flemming, & Nardulli, 1988; See also Engen & Steen, 2000; Ulmer, 1997). In order to answer this research question, I will turn to a series of important legal contingencies that occurred in the U.S. district courts, such

as the Booker (543 U.S. 220, 2005) and the Gall/Kimbrough decisions (552 U.S. 38, 2007, 552 U.S. 85, 2007). In federal courts where judges are precluded from taking gender into account in their sentencing decisions (18 U.S.C. §994(d)), gender disparity favoring female defendants was not only more pronounced compared to that in state courts (Nagel & Hagan, 1982), but also it was claimed that the gender gap increased in the post-guidelines era (USSC, 2004). Regardless of any potential reasons behind this unexpected increase in gender differentials, this counter-intuitive finding begs a simple question that the increased gender gap following the implementation of the guidelines would decline if judges regained their discretion to freely depart from the guidelines. This hypothetical scenario has finally come true, as the Supreme Court declared in the Booker and Gall/Kimbrough decisions that the federal sentencing guidelines are effectively advisory, thus re-equipping federal judges with sentencing discretion. Against this backdrop, in the second half of the current dissertation, I will first examine whether gender disparity has significantly changed following the Booker and Gall decisions, and whether the potential changes, if any, differ across court communities characterized by two particular ecological contexts, religious and political conservatisms.

Taken together, the goal of this dissertation is to contribute to sentencing literature by offering how gender-based sentencing disparity is contingent on broader contexts beyond the immediate circumstances that prior research has examined. Throughout the dissertation, I will attempt to demonstrate that gender disparity in court outcomes is deeply entrenched in ecological contexts related to gender stereotypes or gender role expectations. More specifically, I will determine to what extent the two particular

ecological contexts—political and religious conservatism—affect gender gaps in sentencing outcomes across federal district courts. To accomplish these goals, I will begin by discussing the relevant theoretical and empirical research in Chapter two, showing where the current status of knowledge stands on gender disparity and ecological contexts of courtroom decision-making especially in federal courts. Chapter three will provide detailed contexts on which this dissertation is based, with specific hypotheses for the two sub-studies. In Chapter four, I will offer an overview of methodologies concerning the common themes across the two studies included, followed by separate description of the measures and analytic strategies. Chapter five will present main findings of the two studies. In Chapter six, I will summarize and discuss implications of this dissertation along with recommendations for future studies.

CHAPTER 2: LITERATURE REVIEW

Sentencing Reforms and the Federal Experiment

The History of the 20th Century Sentencing Reforms

During the early periods of American colonization, criminal punishments closely resembled those established in Europe, consisting of corporal and capital punishment with its main goal of vengeance. It was only after the war of independence that the way of punishing offenders moved away from those punishment schemes and it is generally believed that the main goal of criminal punishment in the U.S. was retribution up until 1870 (Rothman, 1983). Throughout most of the 20th century, however, the American criminal justice system in general and criminal courts in particular were permeated by the rehabilitation idea espoused by the Progressive reformers: reducing crimes by treating the needs of offenders (McKenzie, 2001). The idea of reforming offenders through incarceration was first brought to the attention of policy makers around 1870. Criminal punishment was considered a means of reforming offenders, rather than retribution against the behaviors of offenders (Rothman, 1971). Until the first decade of the 20th century, a total of 21 states passed indeterminate sentencing laws—sentencing frameworks without fixed terms of punishment—and by the early 1960s, all of the states in the U.S. had enacted some form of an indeterminate sentencing scheme (Zalman, 1977). Punishment during this period was imposed based on the needs of offenders, not the severity of crimes committed, as the crime was understood as a moral disease that

offenders could not take control of (Cullen & Johnson, 2011). Because sentence severity was predicated on individual defendants' characteristics, criminal justice officials, especially sentencing judges and parole boards, enjoyed a wide range of discretion; at the heart of the progressive reform was the indeterminate sentence, in which the judge metes out a minimum and maximum sentence based on the needs of the offender and the parole board determines when to release the offender by assessing the progress the offender has made toward rehabilitation in prison (Spohn, 2009).

After more than six decades of experiments on this rehabilitative idea, beginning in the late 1960s, these rehabilitation-based criminal justice policies quickly fell out of favor, leading to a dramatic shift to more conservative policies in the justice system (Allen, 1964). The most damaging criticism stemmed from a series of evaluation works designed to assess the effectiveness of correctional programs, which led to the conclusion that rehabilitation of offenders simply does not work (Lipton, Martinson & Wilkes, 1975; Martinson, 1974). Critics emphasized the ineffectiveness of the correctional programs, and also had a deep concern that the coercive rehabilitation programs were inhumane. Collectively, there was a strong sense of failure of correctional programs in addressing the needs of offenders (Von Hirsch, 1976; Wilson, 1975).

Meanwhile, in the field of sentencing, liberal critics also leveled harsh criticisms against the unbridled discretion of judges (Rothman, 1983). The indeterminate sentencing scheme, once heralded as a panacea for crime control, came under fierce attack. One of the most damaging criticisms concerns the equity issue prevalent in the criminal justice

system, ranging from inter-judge disparity in sentencing practices and to disadvantaged treatment of racial minority defendants (Walker, 1993). Most famously, the work of Judge Marvin Frankel laid a foundation for the sentencing reform movement in the early 1970s and afterwards. In his work, *Lawlessness in Sentencing*, the then-federal district court judge delineated serious sentencing disparities among judges and called for an overhaul of the sentencing system (Frankel, 1972).

Sparked by Judge Frankel's provocative essay, both liberals and conservatives joined in the criticisms of the indeterminate sentencing scheme. Those reformers on the left side of the political spectrum claimed that indeterminate sentencing was arbitrary, lacking consistency and predictability in judicial decision-making. In the same vein, potential racial discrimination in courtroom decision-making also raised a serious concern among the liberals and civil rights activists (Davis, 1969; Walker, 1993). In order to address these issues, supporters from this group advocated reducing judicial discretion, which triggered the discussion on sentencing reform. The philosophical rationale supported by this group was a so-called just deserts approach, which rejects the utilitarian approach of sentencing and fundamentally prevents judges from meting out excessive punishments to criminal defendants. The reformers from this camp claimed that sentencing decisions should be driven by the seriousness of crime or blameworthiness, not the needs or risk of offenders, and the current offense should operate as a limiting factor for exacting the amount of punishment (Von Hirsch, 1976).

In contrast, those who championed the need for effective crime control claimed that the offender rehabilitation programs, which played a central role in indeterminate sentencing, were not successful in reducing crimes, as well captured by the famous “nothing works” doctrine (Martinson, 1974). At the same time, these political conservatives also echoed the sentiment that the leniency given to the defendants under the indeterminate sentencing scheme gave the offenders a wrong impression that they can walk away from crime scot-free. Thus, the reformers charged that sentences should be long enough to produce incapacitation effects as well as deterrent effects (Hart, 1968), which has been collectively referred to as the “tough-on-crime” perspective (Rothman, 1983). Reflecting the conflicts in ideologies and specific directions for reforms, the 1970s also witnessed the advent of several hybrid theories attempting to merge one goal of sentencing with another, despite the fact that the goals were not necessarily compatible (Nagel, 1990). To illustrate, Morris (1981) argued that a just deserts approach could be utilized as a sentencing framework, which sets the boundaries on the maximum and the minimum sentences, and the utilitarian approach, such as rehabilitation or deterrence, could be added to fine tune the final punishment imposed (see also Braithwaite & Petit 1990; Walker, 1991).

Against this backdrop, a host of sentencing reforms were proposed and adopted in the late 1970s and 80s. The process began with the adoption of the California Uniform Determinate Sentencing Law in 1977; this was followed by the enactment of determinate sentencing in three additional states: Maine, Illinois, and Indiana. They abolished parole systems and replaced the indeterminate sentence represented by the minimum and the

maximum sentences with fixed terms of punishment set by the judge (Kramer & Ulmer, 2009; Spohn, 2009). Another popular reform was the adoption of mandatory minimum sentencing laws. Nearly every state instituted mandatory minimums in one form or another, thus rejecting any discretion afforded to judges with almost always increasing sentence severity. Some of the famous examples are New York State's Rockefeller drug enforcement laws and the federal system's drug and firearms mandatory sentencing frameworks, which fueled serious controversy (Tonry, 1996).

Finally, 16 states and the federal system took a different path by instituting sentencing commissions, which promulgated sentencing guidelines. Minnesota was the first state to enact guidelines (in 1977) and Pennsylvania did so a year later. The federal guidelines, unarguably one of the most controversial sentencing schemes, were also introduced in 1987 (Spohn, 2009). What needs to be noted, however, is the fact that the sentencing guidelines have taken on quite distinct features depending on the main objectives they wish to achieve. To illustrate, the goals of the sentencing reform movement could be summarized as increasing equity by curtailing judicial discretion and/or escalating punitiveness by raising the severity of punishment. The Minnesota sentencing guidelines placed their main focus on the former, while the Pennsylvania system adhered to both of the goals (Rothman, 1983). The example of the federal sentencing guidelines was an extreme form of sentencing scheme which emphasized both goals to a maximum level (Tonry, 1987; 1996). Below, I will introduce a brief description of the federal sentencing reform movement to provide overall context of this dissertation.

The Case of Federal Reforms: Focusing on the U.S. Sentencing Guidelines

The federal courts were first developed in 1789 by the First Judiciary Act, with the original 13 states designated as district courts. Following the expansion of the U.S. territories and the additional designations of the U.S. district courts, the Circuit Court Act of 1891 formally declared that the federal districts were a primary trial court in the federal system, thus establishing the current judicial system for prosecution of federal crimes (See Federal Judicial Center, 2010 for more information). As noted earlier, federal sentencing had for many years operated on indeterminate sentencing premised on the offender rehabilitation ideal before the nationwide sentencing reform movement overhauled federal sentencing in the late 1980s, which eventually led to the determinate sentencing scheme (Tonry, 1996). Largely responsible for ushering in the determinate sentencing framework in the federal jurisdiction was the Sentencing Reform Act (hereafter “SRA”), a constituent of the Comprehensive Crime Control Act of 1984.

The need for sentencing reform in federal court was first raised by the Brown Commission in 1966, followed by a series of attempts to revitalize federal sentencing practices with its final report published in 1971. In 1976, Senator Edward Kennedy introduced a comprehensive sentencing reform bill, which played a critical role in giving birth to the SRA in 1984; in this bill he proposed that sentencing guidelines be established with uniform goals and that parole be abolished (Nagel, 1990). Against this backdrop, the federal sentencing guidelines—the most controversial and disliked sentencing reform initiative in U.S. history (Tonry, 1996, p. 72) —were formally enacted

on November 1, 1987, following the creation of the United States Sentencing Commission (hereafter “USSC”), by the SRA. The SRA officially recognized three problems of sentencing practices in federal courts during the rehabilitation era, such as disparity, dishonesty, and excessive leniency, with the disparity indicating different sentences for similarly situated offenders, the dishonesty meaning a dramatic disjuncture between the sentence imposed and the sentence served, and the leniency represented as the gap between sentences meted out by judges and the public estimates of how severe sentences should be (Nagel, 1990).

In attempting to address these issues, the SRA set forth some guiding principles on federal sentencing, upon which the USSC was authorized to build the sentencing guidelines. Some of the important directives articulated in the SRA are as follows: first, with regard to the directives on an overall punishment scheme, the SRA specified that the guidelines were to reflect the inappropriateness of imposing imprisonment for the purpose of rehabilitation, providing education or vocational training, or providing medical care or other correctional treatment (18 U.S.C. §994(k)) and the guidelines were to correct the fact that current federal sentences often did not accurately reflect the seriousness of the offense (18 U.S.C. §994(m)). In addition, it was also specified that the guidelines were to determine whether, after conviction, the court should impose a fine, a sentence of probation, or a term of imprisonment (18 U.S.C. §994(a)(1)), and a maximum term of imprisonment was not to exceed the minimum term by more than twenty-five percent or six months, whichever was greater (18 U.S.C. §994(b)(2)).

In establishing offense categories, the Commission was to take into account the following, but only to the extent relevant: the grade of the offense, circumstances of aggravation or mitigation, the nature and degree of harm caused by the offense, community views on the gravity of the offense, public concern generated by the offense, the deterrent effect of a particular sentence, and the current incidence of the offense in the community and nation as a whole (18 U.S.C. §994(c)). Turning to offender categories, the Commission was to take into account the following, but only to the extent relevant: age, education, vocational skills, mental or emotional state, physical dependence, employment record, family ties and responsibilities, community ties, role in the offense, criminal history, and degree of dependence on criminal activity (18 U.S.C. §994(d)). However, the SRA made it clear that the guidelines must be neutral as to the race, sex, national origin, creed, and socioeconomic status of the offender (18 U.S.C. §994(d)); and furthermore, the SRA required that the guidelines, when recommending a term of imprisonment, reflect the general inappropriateness of considering the educational attainment, vocational skills, employment record, family ties and responsibilities, and community ties of the individual defendants (18 U.S.C. §994(e)). Interestingly, the SRA contained some directives in relation to flexibility, and penal resources, as well. For instance, it was directed that the guidelines sentences were to be both certain and fair, while at the same time maintaining sufficient flexibility to permit individualized sentencing when there were circumstances not properly taken into account by the guidelines (18 U.S.C. §994(f)); the guidelines were to take into account the nature and

capacity of penal, correctional, and other facilities and services available (18 U.S.C. §994(g)).

Through a series of discussions and debates which produced multiple drafts along with public comments over the 16-month period, the USSC transmitted the final draft of guidelines to Congress in April, 1987 (Nagel, 1990). A review of historical contexts involving federal sentencing illustrates that federal guidelines are the result of both the just-deserts and the crime control model, thus producing longer but equitable sentences by severely constraining judicial discretion (Tonry, 1987). In determining what kinds of combination of offense and offender categories would result in what level of punishment, the USSC followed what other state sentencing commissions came up with, in which an intersection of offense severity and criminal history scores dictates the level of punishment. Yet, what sets apart the federal guidelines from their state counterparts is the extensively long number of offense levels—the forty-three offense levels, along with a six-category criminal history score (See Figure-2 below). Indeed, the federal guidelines’ extensive classification of offense severity, which determines the base offense level, well contrasts with the 10-grid system of the Minnesota guidelines and the 12-grid scheme of the Pennsylvania guidelines, which also covers misdemeanors. This resulted in what Blumstein and his colleagues (1989, p. 159) called a “sentencing machine” problem. Federal judges considered the extreme form of offense classification system arbitrary and mechanical, and thought that they were being alienated from sentencing (Tonry, 1996).

Figure-2 Federal Sentencing Table

SENTENCING TABLE
(in months of imprisonment)

Offense Level	Criminal History Category (Criminal History Points)					
	I (0 or 1)	II (2 or 3)	III (4, 5, 6)	IV (7, 8, 9)	V (10, 11, 12)	VI (13 or more)
1	0-6	0-6	0-6	0-6	0-6	0-6
2	0-6	0-6	0-6	0-6	0-6	1-7
3	0-6	0-6	0-6	0-6	2-8	3-9
4	0-6	0-6	0-6	2-8	4-10	6-12
Zone A 5	0-6	0-6	1-7	4-10	6-12	9-15
6	0-6	1-7	2-8	6-12	9-15	12-18
7	0-6	2-8	4-10	8-14	12-18	15-21
8	0-6	4-10	6-12	10-16	15-21	18-24
Zone B 9	4-10	6-12	8-14	12-18	18-24	21-27
10	6-12	8-14	10-16	15-21	21-27	24-30
11	8-14	10-16	12-18	18-24	24-30	27-33
Zone C 12	10-16	12-18	15-21	21-27	27-33	30-37
13	12-18	15-21	18-24	24-30	30-37	33-41
14	15-21	18-24	21-27	27-33	33-41	37-46
15	18-24	21-27	24-30	30-37	37-46	41-51
16	21-27	24-30	27-33	33-41	41-51	46-57
17	24-30	27-33	30-37	37-46	46-57	51-63
18	27-33	30-37	33-41	41-51	51-63	57-71
19	30-37	33-41	37-46	46-57	57-71	63-78
20	33-41	37-46	41-51	51-63	63-78	70-87
21	37-46	41-51	46-57	57-71	70-87	77-96
22	41-51	46-57	51-63	63-78	77-96	84-105
23	46-57	51-63	57-71	70-87	84-105	92-115
24	51-63	57-71	63-78	77-96	92-115	100-125
25	57-71	63-78	70-87	84-105	100-125	110-137
26	63-78	70-87	78-97	92-115	110-137	120-150
27	70-87	78-97	87-108	100-125	120-150	130-162
Zone D 28	78-97	87-108	97-121	110-137	130-162	140-175
29	87-108	97-121	108-135	121-151	140-175	151-188
30	97-121	108-135	121-151	135-168	151-188	168-210
31	108-135	121-151	135-168	151-188	168-210	188-235
32	121-151	135-168	151-188	168-210	188-235	210-262
33	135-168	151-188	168-210	188-235	210-262	235-293
34	151-188	168-210	188-235	210-262	235-293	262-327
35	168-210	188-235	210-262	235-293	262-327	292-365
36	188-235	210-262	235-293	262-327	292-365	324-405
37	210-262	235-293	262-327	292-365	324-405	360-life
38	235-293	262-327	292-365	324-405	360-life	360-life
39	262-327	292-365	324-405	360-life	360-life	360-life
40	292-365	324-405	360-life	360-life	360-life	360-life
41	324-405	360-life	360-life	360-life	360-life	360-life
42	360-life	360-life	360-life	360-life	360-life	360-life
43	life	life	life	life	life	life

November 1, 2014

Source: http://www.ussc.gov/sites/default/files/pdf/guidelines-manual/2014/Fine_Table.pdf

Even though the federal guidelines are criticized for the degree to which they constrain judicial discretion, it is also important to point out that the architects of the guidelines did not attempt to totally eliminate judicial discretion in federal courts. Federal judges in the post-reform era retained a small amount of discretion in determining what the appropriate penalty should be. Firstly, depending on the types of behaviors exhibited, the judges are allowed to adjust the base offense level of a defendant. For instance, the defendants could be rewarded for their good behaviors, such as a two-level reduction for acceptance of responsibility (§3E.1.1) and a four-level downward adjustment for playing a minimal role in the offense (§3B.1.2). Similarly, under some aggravating circumstances, the defendants are treated disadvantageously: a two-level increase for crimes committed against a vulnerable victim (§3A.1.1) and a four-level increase for being a leader of a criminal activity that involves five or more participants (§3B.1.1).

However, as these adjustments must be made based on the specific clauses of guidelines and render each detailed element of sentencing decisions visible, critics consider them more as restricting rather than allowing discretion, producing the criticism of the sentencing machine (Nagel & Johnson, 1994).

Secondly, federal judges are also authorized to depart both upward and downward from the guidelines under some circumstances; however, judges must provide justifications in writing to explain the sentences should fall outside of the guidelines (§5K.1.1~§5K.3.1). As will be shown below in greater detail, the discretion to depart from the guidelines has been subject to appellate review and was under tight control,

especially during an earlier period of guidelines implementation. As claimed elsewhere, the grounds for departures were limited and thus the discretion allowed to the judges was nominal (Tonry, 1996), as the guidelines stipulated that some potential grounds for departures were not ordinarily relevant in determining whether a sentence should be outside the applicable guideline range (§5H1.1-§5H1.6); this included factors such as age, education, vocational skills, mental and emotional conditions, physical conditions including drug dependence and alcohol abuse, employment status, family ties and responsibilities, and community ties.

Thirdly, unarguably the most controversial aspect of the federal guidelines is that they required what is referred to as *real offense sentencing* (Tonry, 1996). The USSC invented this contentious sentencing scheme premised on the concern that judicial discretion might be transferred to other criminal justice actors, most prominently to prosecutors (Miethe, 1987). A review of the literature suggests that the commission started assessing the relative advantages and disadvantages of real offense versus conviction-based sentencing systems, finally electing to go with a real offense sentencing system in which the base offense level for the crime is determined by the offense of conviction, which could be modified by other real offense elements (Wilkins & Steer, 1990).

To put it differently, judges were allowed to take into account additional information included in the charges dropped or not filed in modifying the base offense level. Especially problematic is the crime involving the quantity or amount of money,

such as drug trafficking and fraud, because the guidelines factors are often determined by the relevant conduct but, at the same time, mainly drive sentence severity, given the way in which the guidelines were written (Yellen, 2000). Reflecting these deep concerns, this real offense system, which triggered controversy regarding the constitutionality of the guidelines (see Breyer, 1988; Stith, 2008; Tonry, 1996), ended up providing a basis over which the federal guidelines were declared unconstitutional in the landmark decision *U.S. v. Booker* (USSC, 2012).

A second criticism of the federal guidelines, in addition to their extreme rigidity, is that the sentences required under the guidelines were unreasonably severe (Nagel & Schulhofer, 1992; Tonry, 1993). The root of the problem stems from the SRA's stance on sentence severity during the pre-reform era; the SRA stipulated that "the guidelines were to correct the fact that current federal sentences often did not accurately reflect the seriousness of the offense" (18 U.S.C. §994(m)). Reflecting this concern, the commission ignored the specific directive by Congress on the guidelines' consideration of correctional resources (18 U.S.C. §994(g)) and the presumption against the incarceration of non-violent, first-time offenders (18 U.S.C. §994(j)) in devising the guidelines. The USSC achieved its goal of increasing sentence severity by substantially decreasing the use of probation, increasing the average sentence length imposed, and incorporating the mandatory minimums directly into the guidelines (Nagel, 1990). In particular, there was (and is) general agreement that the federal guidelines are overly harsh towards drug

offenders (Nagel & Johnson, 1994; Tonry, 1996).¹ The overall sentiment of the guidelines' extreme severity was also well confirmed by a 2006 survey of federal judges, indicating that a substantial portion of respondents agreed that the federal guidelines, especially for drug offenders, were too severe (Ulmer & Light, 2010). The deep concern toward the excessive punitiveness toward drug offenders led the USSC to vote to apply reduced federal drug penalties retroactively to 46,000 people, reducing average prison terms by 2 years in 2014.

Due to these controversies, the federal guidelines have gone through a series of legal contingencies since their inception. These legal developments well reflect the criticisms directed toward the guidelines, such as constitutionality, rigidity, and severity issues. To begin with, during the earliest period of guidelines implementation, the constitutionality of the guidelines was seriously debated. One of the first challenges that the commission took was *Mistretta v. U.S.* (488 U.S. 361 1989), in which the petitioner Mistretta challenged the constitutionality of the guidelines on multiple grounds, ranging from delegation/separation of power to real offense sentencing. However, the Supreme Court rejected each and every challenge, finally ruling that the federal guidelines were constitutional (Nagel, 1990). Even with the Supreme Court's decision on the legitimacy

¹ This excessive sentence severity toward drug offenders is deeply rooted in the way in which the guidelines are structured (Schuhofer, 1992): first, base offense levels for drug crimes are simply higher as compared to those of other offense types, which is also even more severe relative to sentencing practices during the pre-guidelines era. Second, the substantial increase of base offense level was accompanied by the quantity driven sentencing, which is claimed to be linked to a dramatic increase in punitiveness toward drug offenders. Finally, the way in which the guidelines treat mandatory minimums exacerbated the severity issues by directly incorporating the mandatory minimum into the guidelines structures.

of the guidelines, there still remained some important issues to be resolved, one of which concerned the level of discretion resting with the judiciaries in imposing sentences falling outside of the guidelines. Commentators claimed that, from 1989 to 1996, judges' discretion was held tightly under control, as the guidelines intended, and some even suggested that the discretion was shifted to the prosecutor (Stith & Cabranes, 1998). However, it was *Koon v. U.S.* (518 U.S. 81, 99-100 1996) that reshaped federal sentencing, significantly expanding judicial discretion. In the *Koon* decision, the Supreme Court ruled that an abuse of discretion should be used as a standard for appellate reviews of departures from the guidelines, which resulted in a sharp increase in downward departures awarded by the judges (Stith, 2008). In response to this transition, Congress attempted to counter the increase in downward departures by enacting what is known as the PROTECT (Prosecutorial Remedies and Other Tools to end the Exploitation of Children Today) Act of 2003. This Act, originally aimed at increasing the punishment severity for crimes involving child victims, also restricted judicial discretion for departures, replacing the abuse of discretion standard for departures with *De novo* appellate review of sentences² (USSC, 2006). Following the implementation of the

² The discussion on the standards of review for departure is warranted. To start off with the definition of the standard of review, it basically means the amount of deference given to the decision made by a lower court in reviewing the decision by an upper court. For instance, the abuse of discretion standard, sometimes referred to as the reasonableness standard, assumes a higher level of deference awarded to the original decision. In contrast, the *De novo* standard is the opposite to the abuse of discretion standard. The lowest level of deference is assumed such that the case under review is treated as the case on which decision is made for the first time, when reviewing the case under the *De novo* standard (See Fischman & Schanzenbach, 2011 for more information).

PROTECT Act, judicial discretion was once again argued to be under control and remained at that level before the landmark decision in *Booker v. U.S.* (543 U.S. 220, 2005).

Moving on to the second half of the guidelines implementation era, in the *Booker v. U.S.* decision, the Supreme Court ruled that the federal guidelines premised on the real offense conduct violated the defendants' right to a jury trial and thus declared that the guidelines were advisory. The *Booker* decision introduced the reasonableness of the sentence as the standard of review, which undid what Congress did by enacting the PROTECT Act (Stith, 2008). Following *Booker*, the Supreme Court held that, in the *Gall* decision (128 S. Ct. 586, 2007), judges may not automatically presume the guideline range to be reasonable, and must make an individualized assessment based on the facts presented. In addition, in *Kimbrough v. U.S.* (128 S. Ct. 588, 2007), the court further ruled that, in crack cocaine cases, judges could reasonably conclude that guidelines were not reasonable in an individual case (USSC 2006; 2010). There exists by now a body of research attempting to tease out the impact of the *Booker* and *Gall* decisions on the sentencing practices in federal courts, producing mixed evidence. A group of criminologists claimed that the post-*Booker* sentencing is better captured by the word of stability (Ulmer et al. 2011a; 2011b), while the other camp, led by legal scholars, argued for the instability of sentencing following the landmark decisions (Fishman & Schanzenbach, 2011, 2012; Starr & Rehavi, 2013).

Along with the series of Supreme Court decisions on the unconstitutionality of the guidelines and judicial discretion allowed to the federal judges, in a latter part of guidelines implementation, a few attempts were also made to address the severity issue in federal sentencing. In 2007, the Commission recommended to Congress its revisions to the guidelines concerning the controversial treatments of crack and powder cocaine issue. The guidelines created a one hundred to one ratio between crack cocaine and powder cocaine in calculating base offense levels, which significantly disadvantaged African American offenders, as compared to their white counterparts. Congress responded by passing the Fair Sentencing Act of 2010 in which the mandatory minimum punishments for the drug offenders are substantially reduced (USSC, 2012). Most recently, the Commission also elected to apply a retroactive reduction in punishments among those drug offenders in federal jurisdiction, reducing average prison terms by 2 years in 2014.

In sum, a review of literature and historical contingencies concerning the federal sentencing guidelines suggests that federal sentencing reform was a part of the bigger reform movement which occurred during the 1970s and 1980s nationwide. At the same time, however, it is also clearly shown that the federal experiment has been quite at odds with other states' experiences in many ways, thus creating controversies which were followed by a series of legal contingencies. Against these backdrops, I will turn to the discussion on the gender disparity in sentencing and the ecological contexts of courtroom decision-making, which are the main focus of this dissertation. I will begin with the issue of gender disparity in sentencing.

Gender Disparity in Sentencing

Historical Contexts

Although it might appear as an established fact by now that women offenders are treated with greater leniency over their male counterparts, a close look at the history of social control for female offenders would reveal that, despite the dominant trend of the leniency toward women, this trend fluctuated over time and sometimes exhibited the exact opposite pattern (Boritch, 1992). Tracing back to its origin, it is during the Middle Ages that the notion of chivalry came about; that is, the notion women should be dealt with honor and gentleness (Nagel & Hagan, 1982). More importantly, however, there were also some cases where women were viewed as more immoral and thus were treated with greater harshness compared to their male counterparts. Anecdotal evidence suggests that although the numbers were small, some women were treated more harshly compared to their male counterparts, especially when these women violated the religious beliefs of the time. In addition, there was a time when older women were also more disadvantaged in criminal punishment, as this group of offenders was considered to have lost their reproductive capacity and was oftentimes abandoned by their husbands (Boritch, 1992).

The Progressive Era also marked an interesting shift in the treatment of female offenders (Rafter, 1983). At the heart of the changes in the treatment of women offenders is the shift in the sexual stratification of the labor force, which threatened the prevailing middle-class ideology of separate spheres—men as bread winners and women as care-takers. Reflecting these changes, women offenders who broke the middle-class standard

for women were likely to be punished more harshly during this period. Indeed, a body of research documented that women who committed prostitution or vagrancy faced harsh sentences (Boritch, 1992) and were exposed to distinct formal social mechanisms, such as reformatories (Rafter, 1983), as these women were viewed as immoral and threatening the existing social order (Harris & Firestone, 1998).

Turning to the sex-based disparity during the contemporary era, the traditional take on gender difference in sentencing was that judges would impose more lenient punishment for females relative to their male counterparts, as the purpose of sentencing was premised on utilitarian principles, including rehabilitation and deterrence. Specifically, female offenders were considered a lower risk to the community (Allen, 1987), were seen as having more potential for rehabilitation (Kruttschnitt, 1982) and as more vulnerable to the pains of imprisonment and the effect resulting from labeling processes (Schur, 1984). Another factor justifying more lenient treatment was that women were more likely than men to be responsible for taking care of children (Daly, 1987a, 1987b). However, this traditional view was challenged by the sentencing reform movement in the 1960s and 1970s, which shifted the discourse from the special treatment of women offenders to the notion of gender neutrality in sentencing. In other words, the more recent policy discourse on gender and sentencing suggested that men and women should be treated the same as long as they are similarly situated (Williams, 1982).

Meanwhile, the discussion on gender equality in sentencing did not attract as much attention as race disparity did during the pre-reform period, as observers did not consider

gender disparity to be as problematic as racial disparity (Nagel & Hagan, 1983). However, the advent of the women's rights movement and the emergence of feminist scholarship reshaped the discussion on gender disparity in sentencing, which contributed to the rise of the argument for gender equality in courtroom decision-making (Roberts, 1994; Steffensmeier, 1980). With the participation of feminist legal scholars, a complex philosophical debate followed on whether women should be treated the same as men in courtrooms (see Nagel & Johnson, 1994).

The camp campaigning for the equal treatment argued that the lenient treatment previously accorded female offenders largely stemmed from the negative stereotypes depicting women as weak, passive, submissive, and in need of special protection from men (Williams, 1982). Similarly, Moulds (1980) also claimed that paternalism, one of the theses accounting for the female sentencing discount, equates women with children, thus viewing women as less culpable for their behaviors. Collectively, these scholars feared that the preferential treatment might help perpetuate the negative societal stereotypes of women, making a case that judicial decision making should be gender neutral if males and females defendants are similarly situated (Williams, 1982). In contrast, those scholars from the opposite camp held that, given the biological and cultural difference between men and women, gender and/or gender related factors, such as pregnancy, child rearing responsibility and others, should constitute legitimate considerations for sentencing (Raeder, 1993; Wolgast, 1980). They further maintained that the strict concept of gender equality is neither desirable nor just, as it would actually disadvantage females in the name of equality (Chesney-Lind & Pollock, 1995; Nagel & Johnson, 1994).

Reflecting the complexity and the lack of consensus underlying the gender equality argument in courtroom decision-making, the basic stances on gender equality instituted in sentencing guidelines vary to a certain degree (see Griffin & Wooldridge, 2006). To begin, gender is formally prohibited from consideration when judges mete out the sentences in federal district courts, while the states, such as Pennsylvania and others, do not have the corresponding clause prohibiting judges from taking gender into consideration in their decision-making. In addition, some guidelines, for instance, the federal sentencing guidelines, took an even more restrictive stance on gender equality by specifying that so-called gender-related factors, such as family ties and childrearing responsibility, are not ordinarily relevant in judges' departure decisions (Spohn, 2009). As will be seen throughout this dissertation, the very difference in the view on this issue—that is, to what extent women should be treated differently or similarly relative to their male counterparts—has a lot to do with the way in which people see gender stereotypes and approach crimes committed by women.

Current Evidence on Gender Equality in Sentencing

Next, I turn to the current status of knowledge on gender equality in courtroom decision-making. Gender disparity has been relatively well documented in sentencing literature (see Bickle & Peterson 1991; Bontrager, Barrick, & Stupi, 2013; Daly & Bordt, 1995; Nagel & Hagan, 1983). Overall, there exists a consistent pattern of favoring female defendants over male defendants in both state courts (Spohn, 1998; Spohn & Spears, 1997; Steffensmeier et al., 1993; Steffensmeier et al., 1998) and federal courts (Albonetti,

1997; 2002; Brennan & Spohn, 2009; Mustard, 2001; Stacey & Spohn, 2006). Moreover, this pattern is found in both the pre-sentencing reform era (Hagan, Nagel, & Albonetti, 1980; Nagel, Cardascia, & Ross, 1980) and the post-sentencing reform era (Starr, 2012; Steffensmeier & Demuth, 2006), though a varying degree of heterogeneity is also reported (Daly & Tonry, 1997; Spohn & Beichner, 2000). A review of the literature also reveals that this pattern of preferential treatment is more pronounced in the incarceration decisions over sentence length decision (see Daly & Bordt, 1995; Nagel & Hagan, 1983, but see Steffensmeier et al., 1993).

Whereas a majority of scholars attributed this observed disparity to real sex effects (Bickle & Paterson, 1991; Rodriguez et al., 2006), some scholars explained it as a statistical artifact resulting from the failure to employ rigorous controls (Kruttschnitt & Green, 1984; Steffensmeier, Kramer, & Streifel, 1993; Wooldredge, 1998), thus rejecting the idea that women were treated more favorably as compared to their male counterparts. For instance, Steffensmeier and his colleagues demonstrated that introducing more refined control variables rendered the gender effect negligible at the in/out decision and null at the sentence length decision (Steffensmeier et al., 1993).

Taking this line of argument one step further, Daly (1994) claimed that it was inevitable to find a sex effect, especially when relying on traditional methods of analyzing statistical data. In particular, she held that comparing like crimes is a myth in estimating gender disparity with the quantitative statistical analysis, and she further maintained that gender disparity in sentencing would be minimal when taking into

account the difference between the seemingly similar crimes. However, even though it is worthwhile to appreciate the challenge regarding the measurement issue concerning gender disparity, evidence from a large body of studies appears to support the claim that the female discounts are real (See Bontrager et al., 2013; Daly & Tonry, 1998; Starr, 2012).

Research also suggests that the gender effect may be conditional on other important factors, such as race, marital and family status, and offense type, independent of its main effects (Spohn, 2009). The underlying theme in these bodies of research is that not all woman offenders enjoy the female discount. Rather, the benefit is applied to certain groups of female offenders. To start, a number of scholars have addressed the issue of race and gender interaction, arguing that chivalry in court sanctioning typically bypasses women of color (Klein & Kress, 1976; Richey-Mann, 1995). Belknap (2001), one of the leading authorities on this topic, argues that chivalry is inherently racist, structured to protect respectable (that is, middle-class and white) women only, and thus minority female offenders, who disproportionately come from disadvantaged backgrounds, will fail to occupy this respectable female role. Previous studies reported equivocal results about this issue. Specifically, whereas some researchers reported race effects favoring white females over Black females (Crawford, 2000; Kruttschnitt, 1982; Steffensmeier et al., 1993, 1998), others uncovered differences that favored Black females (Griffin & Wooldredge, 2006; Koons-Witt, 2002). There are also studies that found that race/ethnicity did not affect sentence severity among women offenders (Bickle & Peterson, 1991; Spohn & Beichner, 2000; Steffensmeier & Demuth, 2006).

Another important body of research regarding the conditional effect of gender involves the interaction between crime types and gender. One line of research in this tradition was to examine the relationship premised on the idea that female offenders who committed more masculine crimes would not be accorded lenient treatment (Moulds, 1980; Rache, 1975). Spohn and Spears (1997) investigated this issue, finding that even those female defendants who committed violent crimes were more likely to have their charges dismissed, less likely to be incarcerated, and more likely to receive shorter sentence over their male counterparts (see also Farnsworth & Teske, 1995; Mustard, 2001). In contrast, Steffensmeier and his colleagues (1993) found that shorter sentences were imposed on females convicted of a serious felony, while relatively longer sentences were meted out to female defendants convicted of less serious crimes (see Daly, 1987; Steffensmeier et al., 1998). Finally, the USSC report (2004) provided evidence against this presumption that female offenders convicted of masculine crimes would fare differently relative to females convicted of more traditional crimes, finding that females were treated leniently both in drug and non-drug crimes in federal district courts. Collectively, even though there existed a handful of studies providing support for the hypothesis, the weight of evidence clearly indicates that the interaction between gender and crime types is not substantial.

Yet another important body of research concerns gender effects in combination with what may be classified as gender-related factors, including marital status, child rearing responsibilities, pregnancy, and others (Bickel & Peterson 1991; Crew, 1991; Daly, 1987; Kaukinen 1995; Stacy & Spohn, 2006). This line of research was predicated

on the assumption that female offenders who were married or who had children would enjoy better sentencing outcomes, as there might be actual social costs involved with incarcerating these types of female offenders, who closely followed the traditional gender roles as a mother or a housewife (Raeder, 1993). There are a number of studies which found this hypothesized relationship. For instance, Daly's research suggests that females are treated less harshly when they have more familial responsibilities (1987a; 1987b; see also Koons-Witt, 2002; Kramer & Ulmer, 1996). However, there also exist some studies reporting the opposite or null findings. To illustrate, Stacey and Spohn (2006), using data on sentences imposed in three federal district courts, examined whether the marital status and child rearing responsibility condition the effect of gender on sentencing outcomes. The results revealed that even though there was a strong independent gender effect, favoring female defendants, there were no significant interactions between gender and these two gender related factors (see also Brennan & Spohn, 2009; Griffin & Wooldredge, 2006; Mustard, 2001).

In summary, a review of the literature suggests that there is strong and convincing evidence for gender effects favoring female defendants in sentencing. In fact, Daly and Bordt (1995, p. 160), concluded that "sex effects are more frequent and greater than race effects," despite some inconsistencies concerning the conditional effects of gender (see also Bontrager et al., 2013). The question, of course, is how to account for the gender disparity. That is, *why* does the female discount exist, which is deeply entrenched in theoretical approaches to understanding the gender effect in courtroom decision-making.

In what follows, I will lay out important theoretical approaches to explain gender disparity in sentencing.

Theoretical Approaches to Gender Disparity in Sentencing

Traditional perspectives on gender disparity are represented by the chivalry/paternalism thesis and the evil women hypothesis (Spohn, 2009). The chivalry perspective argues that women in general are treated more leniently than their male counterparts, and this difference results mainly from chivalrous attitudes on the part of criminal justice officials, who are mostly male (Curran, 1981). Historically, chivalry broadly refers to the customs and practices developed in Europe around the mid-12th century, which involved special treatment for ladies who were considered weak by Knights (Anderson, 1976). Thus, women were perceived as in need of protection in every domain of society, including legal proceedings, and, accordingly, female offenders were treated with greater leniency relative to male offenders. Similarly, paternalism is another theoretical approach by which lenient treatment toward women is invoked and oftentimes equated with the chivalry thesis (Belknap, 2001). According to Moulds (1980), however, paternalism, which equates women with children, views women as lacking capacity for making responsible decisions and thus, as less blameworthy for their criminal behavior.

The opposing hypothesis is the evil woman thesis, which contends that female offenders who violate sex-role stereotypes, or/and commit more masculine violent crimes, are treated either no differently or more harshly than males who commit the same crimes (Farnworth & Horan, 1980). In Puritan society, females were treated as witches, and thus

were punished more harshly than their male counterparts, when they deviated from the gender norms concerning femininity (Karlsen, 1998). In a more recent perspective, Rasche (1975) implied that evil women could be portrayed as career criminals, perceived as a danger to society, and described even as supernatural women or witches. She also noted that very few women were labeled as evil. According to the evil women hypothesis, female offenders would be punished more severely than male offenders because they committed a double offense: one for breaking the law and the other one for violating gender norms (Bontrager et al., 2013). Meanwhile, some scholars acknowledged that the chivalry/paternalism and evil woman hypotheses are complementary rather than competing explanations for the sanctioning of female offenders (Crew, 1991; Nagel & Hagan, 1982). These scholars do not classify chivalry as representing general leniency. Instead, they argued that female offenders who only commit feminine offenses receive lenient treatment. In a similar context, Herzog and Oreg (2008) also offered a renewed distinction between “true chivalry” and “selective chivalry (the evil women hypothesis),” charging that these two hypotheses are premised on the same causal explanation. Conceptualized in this way, sex roles or stereotypes that flow from the different sex roles play a key role in accounting for gender disparity, which led some scholars to refer to these theories as sex roles theory (Kruttschnitt & Savolainen, 2006).

After reviewing studies published before 1982, Nagel and Hagan (1982) concluded that women tended to receive preferential treatment, thus the chivalry thesis seems to have prevailed over the evil woman thesis. But, as Daly (1987) convincingly argued, the intention of judges had never been adequately accounted for, which means

that there is only indirect empirical support for the chivalry/paternalism hypothesis. Reflecting the conceptual issues, some commentators criticized these two traditional approaches, contending that “concepts such as chivalry/paternalism and evil women lack an empirical referent and analytical bite” (Daly & Tonry, 1997, p. 234). More recent perspectives attempted to add more specificity to the explanation of gender disparity in sentencing outcomes.

One of the first attempts could be attributable to a series of studies conducted by Kruttschnitt and her colleagues (Kruttschnitt, 1982; 1984; Kruttschnitt & Green, 1984; Kruttschnitt & McCarthy, 1985). Building on Donald Black’s (1976) theory arguing for an inverse relationship between formal and informal social control, she contended that women offenders are accorded greater leniency due to the higher level of informal social control exerted upon them, relative to their male counterparts. More specifically, Kruttschnitt underscored the importance of women’s dependency both in economic and non-economic forms. As she noted, women were treated more leniently than men in courtroom decision-making because women were more likely to be economically dependent on men and because a high level of supervisory activity was associated with women stemming from residing with others.

Daly’s work also paved an important avenue for understanding gender disparity in sentencing (1987a; 1987b). Her theoretical approach, known as familial paternalism, highlights the importance of the social costs associated with punishing female offenders. This makes an interesting comparison with Kruttschnitt’s take on the importance of social

control in explaining gender gap in sentencing. Daly contends that women are treated more leniently not because there is chivalry involved with punishing female offenders (female paternalism), but because sending women to prisons involves too much social cost, especially when they have family obligations to perform (familial paternalism). More specifically with regard to gender disparity, she further argued that women's role in the family as a care-taker is not easily replaceable, while men's role as a bread winner may be more readily replaced by the welfare system.

Even though these two theoretical frameworks clearly improve upon the traditional chivalry/paternalism theses in many ways, by far the most influential theoretical approach to accounting for extra-legal disparity is the focal concerns perspective advocated by Steffensmeier and his colleagues at Pennsylvania State University (See Kramer & Steffensmeier, 1993; Steffensmeier et al., 1993; Steffensmeier et al., 1998; Ulmer, 1997). Building on Albonetti's uncertainty avoidance and causal attribution perspective (1986; 1987; 1991), they argued that judges' sentencing decisions were mainly guided by concerns about offender blameworthiness, dangerousness, and the practical constraints or social costs of sentencing decisions. However, the lack of complete information on offender culpability and dangerousness leads judges to develop a perceptual shorthand based on societal stereotypes linked to offender characteristics such as race/ethnicity, age, and gender, as these social stereotypes could reduce the inherent uncertainty involved with their sentencing decisions (see also Spohn & Holleran, 2000; Steffensmeier et al. 1998; Steffensmeier & Demuth, 2000).

In their attempt to account for gender disparity in sentencing, Steffensmeier and his colleagues (1993) maintained that judges who sentence female offenders more leniently may be motivated more by the two main focal concerns—blameworthiness and practicality—than by paternalism or by beliefs that women who commit crimes are evil. Blameworthiness is a rather loose concept, which may incorporate a range of different aspects of defendant behaviors (Wheeler, Mann, & Sarat, 1988). Oftentimes it is equated with the terminology of offender culpability, which only represents the seriousness of offending behaviors (Von Hirsch, 1976). However, a more general definition of blameworthiness is premised on a “broader meaning of the harm in light of the defendants’ social history and behavior” (Daly, 1994, p. 174), which refers to “the broader linking of the defendant’s biography (social history and prior record) to the offense” (p. 40). Similarly, Steffensmeier et al. (1998) also made clear that the concept of blameworthiness incorporated not only the culpability of offenders represented by the seriousness of the current offense, but also a variety of offenders’ biographical factors.

It has been claimed that, as compared to their male counterparts, female defendants are considered less blameworthy because of their nonviolent prior records, minor roles in the offense, and remorseful attitudes. For instance, a judge interviewed in Pennsylvania indicated that women and men are apples and oranges in terms of blameworthiness, even when they show similar characteristics on paper. More specifically, the judge went on to say that “a woman coming before you in court may have the same prior record score or the same offense score as a man but her score involves all property offense—no violent priors at all. And many times the woman’s part

in the offense is small, more the follower than the leader.” (Steffensmeier et al., 1993, p. 434) In addition, the history of child abuse and sexual victimization, which oftentimes characterizes the troubled lives of female defendants, also makes the women appear less blameworthy. Similarly, Daly (1994) makes a persuasive argument that the blurred boundary between victimization and offending behaviors for female defendants oftentimes leads the judge to perceive the female defendants with victimization experiences in a more sympathetic light, as compared to their male counterparts with and without the history of victimization.

The other focal concern which may prompt judges to impose more lenient sentences is the practical constraints on and consequences of sentencing decisions. According to Steffensmeier et al.’s (1998) account, this concern is closely related to both organizational constraints and individual defendants’ consequences resulting from judicial decision-making. Pragmatic considerations associated with sending women to prison, such as organizational demands of incarcerating pregnant women or women with physical or mental health problems, and less available jail and prison space for female defendants, may further impel judges to sentence female defendants more leniently (Steffensmeier et al., 1993; Ulmer, 1997). More important concerns may be any potential damaging effects on the children or the families of the female defendants.

This point is well illustrated by a quote from a court-actor interviewed by Ulmer (1997, p. 134): “I represent a young girl right now, 22 years old. She has two children, no husband.....It is a mistake to put a person like that in jail, because of the circumstances.

What will happen to the kids?” There is a body of literature, which examines to what extent so-called familial women received preferential treatment in the courtroom (see Stacey & Spohn, 2006, for example). Most notable is a series of studies done by Kathleen Daly (1987a) in which she contended that judges’ sentencing decisions were mainly guided by their intent to protect children and families, which she refers to as familial paternalism. In her theorizing, it is important to note that having dependents, rather than being married, is the key family status factor leading to more lenient outcomes (Daly, 1987b). She also added that these family circumstances had more pronounced mitigating effects on outcomes for female defendants than for male defendants.

The Importance of Gender Norms in Sex-Based Disparity

The review of the major theoretical approaches so far reveals the important insight that gender norms play a critical role in producing gender differentials in sentencing outcomes: both the chivalry and evil women hypotheses suggest the possibility that the degree to which women offenders deviate from the prevailing gender norms influences the level of criminal punishment (Nagel & Hagan, 1982). Similarly, the focal concerns perspective also highlights the importance of gender stereotypes in accounting for sex-based disparity. As suggested, the two elements of focal concerns, blameworthiness and the practical constraints, are said to be strongly impacted by gender-based stereotypes due to the bounded rationality associated with the lack of complete information on specific cases and specific offenders (Steffensmeier et al., 1998). In spite of such theoretical importance given to gender norms as an underlying cause for explaining

gender disparity in sentencing, surprisingly little attention has been paid to the discussion of the nature of gender norms and even less attention has been devoted to understanding the processes by which gender stereotypes impact judicial decision-making.

Gender norms are a broad set of rules for behaviors and attitudes appropriate for men and women, which are culturally defined. Brannon (2011, p. 160) discusses gender stereotypes and gender roles as follows:

“A gender stereotype consists of beliefs about the psychological traits and characteristics of, as well as the activities appropriate to, men or women. Gender roles are defined by behaviors, but gender stereotypes are beliefs and attitudes about masculinity and femininity. The concepts of gender role and gender stereotype tend to be related. When people associate a pattern of behavior with either women or men, they may overlook individual variations and exceptions and come to believe that the behavior is inevitably associated with one gender but not the other. Therefore, gender roles furnish the material for gender stereotypes.”

The origin of gender norms dates back to the Bible, which describes women as inferior to men (Packer, 1991), but the more contemporary form of gender stereotypes has its roots in the 19th century Victorian era. This period, marked by dramatic societal changes in the wake of the Industrial Revolution, created a sexual stratification of labor—men as a breadwinner working outside the home and women as a caretaker working at home, which in turn produced the doctrine of separate spheres (Lewin, 1984). The cult of true womanhood also emerged around the same time period; this underscored the values of piety, purity, submissiveness, and domesticity among women (Welter, 1978).

The virtue of piety suggested that women were naturally more religious than men. For instance, the Bible, portraying women as more submissive and as a follower, suggests

a male-dominated system both in society and family. The second virtue was purity. It was considered that the loss of purity, often equated with virginity, was as bad as death and thus purity was deemed as essential to a young woman (Welter, 1978). The third virtue concerned submissiveness. According to this principle, women were expected to be weak, dependent, and timid, whereas men were supposed to be strong, wise, and forceful. In a similar way, it was further suggested that a wife should conform to the authority of her husband. Finally, the concept of domesticity implied that women should stay at home, dealing with domestic affairs (Welter, 1978). In a similar context, traditional middle-class ideology of women also indicated that women should conform to the ideals which underscored their moral purity and motherhood (Cott, 1987; Roberts, 1976). As suggested, the separate spheres doctrine confined women to home where females were obligated to give birth and raise children (Lewin, 1984). So there were some conventional female roles expected in marriage and family, which have important ramifications on gender norms.

The traditional ideology of femininity is also related to religious values. At the center of this middle-class gender norm lies the mores of Christianity. According to Roberts (1990, p. 289), “Christianity has a long history of sex bias.” For instance, the Bible, portraying women as submissive and as followers, suggested a male-dominated system both in society and family. Even in contemporary society, there exist negative attitudes toward female clergy, who are treated disadvantageously compared to their male counterparts.

Another important ramification of the gender norms could be found at the intersection of gender, class, and race. Some scholars claim that the prevailing gender norm is based on white, middle-class women; they argue that there are race-specific gender norms (Belknap, 2001). The point of departure for this fine distinction relates to a broader social structure surrounding women of color. As suggested, the traditional concept of femininity is deeply entrenched in the issue of class (Hahn, 1980), as the sexual stratification of labor, best represented as separate spheres, did not apply to lower class African American and Hispanic women (Boritch, 1992). Unlike their white counterparts, women of color were forced to work outside of their homes due to their inexpensive labor costs as servants, factory workers and etc. (Klein & Roberts, 1974). Thus, it is more likely the case that women of color have been perceived as not exhibiting traditional feminine behaviors and virtues.

More specifically, African American women were depicted as independent, as the historical socio-economic structure was deeply rooted in slavery (Gilkens, 1983), which in turn exposed Black women to the plantation life. Similarly, Black women were also described as domineering. Historically the common family structure among Blacks was a female-headed household in which the mother has more authority than the father (Young, 1986). Finally, Black women were also described as highly sexual (Rome, 1998) and as baby-making machines (Fishman, 1998). Their sexuality and motherhood did not deserve the respect that their white counterparts enjoyed throughout the history. Collectively, these images help to defeminize African American women, thus subjecting them to the perception that Black women are more culpable for their behaviors, as compared to white

females (Brennan, 2006). To a lesser degree, there also exist similar but distinct societal stereotypes for Hispanic women. Hispanic women are oftentimes described as gang members or mothers of gang members due to the images associated with drug smuggling (Portillos, 1998). Sexuality also constitutes another important dimension that characterizes Latina women; Hispanic women are likely to be perceived as sexy women of easy virtue (Castro, 1998).

Despite these rich literatures on gender norms, one important question that did not receive much scholarly interest concerns ways by which gender norms influence judicial decision-making. One simple explanation may be that the judges, who are a part of a broader community, share some fundamental characteristics related to gender norms and they automatically rely on their beliefs in their decision-making. Indeed, developmental literature demonstrates that kids as young as three years old could distinguish words and objects gender appropriate and six year- old children already practice gender-based decision-making (Martin & Little, 1990). Researchers suggest that gender-based decision-making is established early in life and thus it is possible that judicial decision-making is influenced by gender stereotypes that the judges endorse. Conversely, a recent development in social psychology literature on stereotypes concerns the issue of implicit cognition. Mainly based on the implicit association test (IAT) method,³ this new line of

³ IAT is one of the most frequently used implicit measures of stereotyping (Greenwood et al., 1998). This method was developed using the concept of an automatic association in order to overcome the weaknesses of the traditional way of measuring stereotypes- a self-report method, such as social desirability bias (See Rudman & Glick, 2001 for more information on IAT).

research finds that oftentimes people do not overtly rely on social stereotypes but rather the use of social stereotypes is implicit (Greenwald, McGhee, & Schwartz, 1998). In this regard, Rudman and Glick (2001) demonstrated that an implicit measure of gender stereotypes successfully predicted job discrimination against women rated as more agentic, while the explicit measure of gender stereotype, based on a self-report method, did not perform well. Nosek et al. (2002) also found that, in their study based on the IAT method, men were more likely to be associated with math and career, and women were more likely to be linked to liberal arts and home. Equally important, there is also an emerging body of literature which suggests that people do not need to personally endorse the social stereotypes to exhibit those stereotypical behaviors (Bargh, Chen, & Burrows, 1996). Collectively, this scientific knowledge points to the possibility that sentencing judges, who likely were trained as professionals not to resort to social stereotypes, could be also influenced subtly or implicitly by these social forces.

Finally, it is important to point out that gender norms are also defined by culture and social structure, which reciprocally influences the social status of women in society. For that reason, gender norms vary across different nations and different regions within a nation. Indeed, a small body of gender literature demonstrates that attitudes toward gender norms significantly differ across regions (Boles & Atkinson, 1986). For instance, Rice and Coates (1995) showed that Southerners are much more likely to hold conservative gender role attitudes, compared to non-Southerners. Then there exists a reason to believe that gender disparity in sentencing outcomes could vary across different ecological contexts, as the gender norms also differ across geographical locations. This

endeavor was proposed almost three decades ago by Nagel and Hagan (1983, p. 136) as one of the promising avenues to advance literature on gender and sentencing. But the effort to uncover this important relation did not come to full fruition until recently. In what follows, I will review theoretical and empirical literature on social ecology of sentencing to lay the foundation for the discussion on gender disparity and social ecology of courtroom decision-making.

The Impact of Ecological Contexts in Sentencing

The idea that the level of criminal punishment may vary across different locations creates a dilemma, as there exists a potential conflict between the principle of equal justice and the notion of local autonomy, both of which underlie the fundamental values of Americans (Ulmer & Johnson, 2004). Similarly, the main goal of the sentencing reform movement—promoting consistency and uniformity in judicial decision-making, thus reducing both between and within jurisdictional idiosyncrasies—largely collides with the community-based movement, which underscores the importance of community as a collective entity and emphasizes each community’s unique local norms and resources (Lubitz & Ross, 2001). Reflecting this complexity, relatively less attention was devoted to reducing inter-jurisdictional differences in sentencing practices than decreasing inter-judges’ disparities within a jurisdiction, even though the inter-jurisdictional discrepancy in sentencing practice was clearly recognized as one of the impetuses behind the sentencing reform movement at the federal level (Breyer, 1988).

Theoretical Frameworks: Focusing on Court Community Perspective

In the sociological literature, there are theoretical attempts to account for the reasons behind the idea that criminal punishment could vary across ecological contexts. To begin, classic literature on the sociology of law suggests that legal punishments vary across communities as the level of formal social control is shaped by the social organizations and collective sentiments of communities (see Durkheim, 1966; Ehrlich, 2002). More recent literature on social control also echoes the idea that a variety of social arrangements and social climates affect the level of punishment applied in the community (see Garland, 2001; Savelsberg, 1994; Tonry, 2009). In addition, there also exists a long line of research premised on the conflict perspective, maintaining that punishment processes are inherently political, as the criminal justice system is devised as a way to control the underclass and to perpetuate the social structure along the class line (Chambliss, 1989; Sampson & Laub, 1993).

Reflecting the fact that social environments affect judicial decision-making, a recent body of literature in sentencing, referred to as the contextual perspective, claims that judicial decision-making is shaped by courts' environmental factors as well as individual case characteristics, (Helm & Jacobs, 2002; Myers & Talarico, 1987; Peterson & Hagan, 1984; Sampson & Laub, 1993). To illustrate, in their seminal work, *The Social Contexts of Criminal Sentencing*, Myers and Talarico (1987) make a persuasive case that the findings of previous research on sentencing should be reinterpreted from the contextual angle, placing an importance on the ecological and temporal arrangements of

court environments. This new theoretical approach to sentencing notes the importance of understanding multiple sources of variation in sentencing outcomes, which goes beyond individual case level characteristics (see also McIntosh, 1991). One of the most theoretically sophisticated attempts designed to unpack between-court variations in sentencing is the court community perspective and largely responsible for developing and fine tuning the perspective is a series of studies conducted by Eisenstein and his colleagues (Eisenstein, Flemming, & Nardulli, 1988; Eisenstein & Jacob, 1977; Flemming, Nardulli, & Eisenstein, 1992; see also Ulmer, 1997). Based on their field research in the felony courts in Pennsylvania, Michigan, and Illinois, they discovered that the county courts differed substantially in the ways in which the courtroom work groups processed cases, which constitutes the core hypothesis derived from the court community perspective. The authors conceptualized each court as a court community with its own distinctive culture and organizational arrangements. They maintained that local court communities promote their own local court culture or substantive rationalities and these local cultures in turn shape sentencing processes and outcomes as much as formal policies and legal structures such as sentencing guidelines (Eisenstein et al., 1988).

In order to better appreciate the theory, it is important to have a good grasp of the two important concepts subsumed under the perspective. An overarching concept is the concept of the court as a community. The authors claimed that conceiving the court as a community allows for incorporating not only the inner dynamics of courtroom workgroups, but also the influences of external environments to the court organization. According to the authors' exposition, the court community is characterized by two

important structural features: a common workplace and interdependence among workgroup members. The former element concerns the inhabitants of the court community, basically suggesting that court actors working at the courthouse, usually characterized by formal membership in the sponsoring agencies, are the inhabitants of a court community. In this regard, it is also noteworthy to consider Ulmer's (1997) similar approach, which looks at the court as a social world. Ulmer views the court as a social arena in which different court actors interact to attain common goals. What sets this approach apart from the court community perspective is the fact that the boundaries of the court community are set by the lines of communication and participation around shared tasks, not by the formal organizational membership. With regard to the latter component—interdependence—it was argued that courtroom actors are enmeshed in a complex web of relationships, whose interactions substantially influence each other, finally producing variation in how courtroom workgroups work across different court communities.

Another important concept within the court community perspective is local legal culture, which is defined as “the values and perceptions of the principal members of the court community about how they ought to behave and their beliefs about how they actually do behave in performing their duties” (Einstein et al., 1988, p. 28). The authors provide further clarification on what characterizes the local legal culture. They suggest that for something to qualify as an indicator of local legal culture there must be some level of consensus and stability. More specifically, the authors argue that local legal culture includes such things as a special language and non-verbal expressions only

available in the community and the shared beliefs about how cases are (and should be) processed and about the nature of interpersonal relationships. At the heart of local legal culture—at least from a sentencing perspective—lies the concept of going rates for punishment. Going rates refer to informal sentencing norms about the appropriate punishment for certain kinds of crime and certain kinds of offenders and offer a kind of template by which the courtroom workgroup members process their cases in an efficient way (Einstein & Jacob, 1977; Einstein et al., 1988). Einstein and Jacob (1977) argued that going rates set the boundaries on the interaction among courtroom workgroup members, thus making case processing as efficient as possible with a lower level of uncertainty involved with case processing.

Another key prediction from the court community perspective is related to the impact of changes in formal norms or sentencing policy on sentencing practices. The traditional perspective is premised on the assumption that any changes in sentencing policies would directly translate into the changes in sentencing practices (Lukacs, 1971). This position, oftentimes referred to as the formal legal model, suggests that judicial decision-making revolves around legally relevant variables and thus changes in sentencing policies—formal sentencing norms—directly lead to changes in the sentencing practices of individual judges. An opposing view comes from Savelsberg's (1992) concepts of substantive and formal rationality; he claimed that sentencing reform movements in the 1970s and 1980s would fail because formal rationality, represented by the sentencing guidelines, cannot defeat the substantive rationality developed within the court community and reflecting unique needs of court communities.

As a slightly different perspective, the court community perspective holds that the intended goals of policy changes are always filtered through the informal norms of the court community, thus leading to heterogeneous outcomes across different court contexts (Eisenstein et al., 1988). This idea of filtering is well established in sociological literature on inter-organizational relations, in which externally imposed rules are always modified by organizational cultures and structure (Fine, 1984). In a similar context, Ulmer and Kramer (1998) borrowed the concept of embeddedness from network theory (Granovetter, 1985), arguing that an everyday use of sentencing guidelines is embedded in local court contextual factors. Indeed, Ulmer (1997) systemically demonstrated that the implementation of sentencing guidelines in Pennsylvania varied significantly across court communities, depending on the level of informal sentencing norms: some courts with weak county culture were more likely to be responsive to the intention of the sentencing commission, while other courts with strong going rates did not embrace the changes in policies. Reflecting this variation in the implementation of sentencing guidelines, Ulmer claimed that “the nature and character of ... formal social control ... depend as much or more on the processual orders of local courts as they do on the policies” (p. 189). And it was further maintained that the differential implementation of sentencing guidelines was due to the degree to which the sentencing reform movement was in line with the local order of the court community.⁴

⁴ Embedded within this line of research is the work of Engen and Steen (2000). Their study assessed the impact of guideline changes for drug offenders in the state of Washington. The study demonstrated that the changes in the guidelines led to the intended changes in sentencing outcomes, thus providing support for the argument made

Meanwhile, it is also worthwhile to note that the court community perspective is a comprehensive theoretical framework within which other individual and group level theories in sentencing literature could be situated (Kramer & Ulmer, 2009). To illustrate, the focal concerns theory is enmeshed with the court community perspective in many ways: the interpretation and prioritization of focal concerns is assumed to vary between courts, because they are embedded in local court communities' organizational and cultural milieus (Ulmer & Johnson, 2004). At the group level, the characteristics of courtroom workgroups are said to shape local legal culture, which in turn influences the way courtroom workgroup members interact with each other (Eisenstein & Jacob, 1977). At the macro level, the court community perspective is also congruous with diverse sets of theoretical explanations developed within structural and organizational approaches (Ulmer, 2012). In its relation to the organizational efficiency perspective, which highlights the importance of efficient case processing by inducing guilty pleas as an organizational goal (Dixon, 1995), the court community perspective, which emphasizes the role of informal sentencing norms, shares the fact that the processes in relation to plea bargaining are strongly governed by the local legal culture in the court community (Eisenstein & Jacob, 1977; Eisenstein et al., 1988). At the macro level, the social threat perspective could be also integrated with the court-community perspective (Ulmer &

from the formal legal model of sentencing. However, this study also found that the resulting changes were not universal; rather, they were conditional on case processing factors. To illustrate, it was found that the increased sentence severity for drug offenders—the desired goal of the guideline modification in Washington—was bigger for those defendants who went to trial. The increased penalty was smaller for those defendants who pled guilty, assumed to be more strongly regulated by informal sanction norms, rather than formal sanction norms, such as the policy changes.

Johnson, 2004). The racial threat perspective, one of the dominant theoretical approaches among others, states that the size of racial and ethnic minority population increases punishment severity (Britt, 2000; Spohn, 2000), as Blacks and Hispanics in American society are more likely to be perceived as objects of fear and thus more threatening than their white counterparts (Blumer, 1955). Local legal culture, the most important concept derived from the court community perspective, is argued to be influenced by community-level structural factors, one of which involves racial/ethnic composition of the community. Thus, the level of perceived fear of and perceptions of the dangerousness of racial/ethnic minorities is likely to be reflected in the legal culture. And, at the same time, it was also suggested that racially charged actions and beliefs often become a part of case processing strategies embedded within the normal operation of institutions (Bonilla-Silva, 1997).

To summarize, a review of the court community perspective suggests that judicial decision-making is a complex process upon which diverse factors from multiple layers of social environments exert influences. Of particular importance is the idea that social arrangements and climates, once not considered critical, play an important role in shaping sentencing outcomes. The essence of this notion is well capsulated by a quote from Ulmer and Johnson (2004, p. 137), who are the two leading experts on this topic: “both the level of and the criteria for punishing criminal defendants vary ... from place to place.” Finally, it was also suggested that the implementation of sentencing reforms would also have a distinct impact based on the characteristics of court community, which both directly and indirectly shapes court outcomes. In what follows, I will review relevant

literature on community characteristics that may explain the community-level variation in court outcomes.

Empirical Evidence of Key Contextual Variables

As explained above, one of the key predictions made from the contextual approach in sentencing is that there will be significant jurisdictional variation in sentencing outcomes, net of individual- and case-level characteristics. The current empirical literature lends support to this argument by documenting that not only the odds of incarceration (Ulmer & Johnson, 2004; Wang & Mears, 2010a, 2010b; Weidner & Frase, 2003), but also the average sentence lengths (Myers & Talarico, 1987; Ulmer & Johnson, 2004) significantly vary across different court communities. Meanwhile, despite the well-established fact that the application of criminal sanctions differs across jurisdictions, it should be also noted that the contextual variables typically do not account for a substantial portion of variance in individual sentencing outcomes. In this regard, Bushway and Forst (2013) describe the utility of these contextual variables in explaining between-court variations as disappointing. One particular study, however, needs to be recognized with regard to the issue of the limited explanatory power of contextual variables. Baumer and Martin (2013) pointed out that extant empirical literature on ecological contexts and sentencing has paid almost exclusive attention to what they called structural conditions of community and indirect measures of key variables (e. g., % Black as a proxy for perceived racial threats), thus yielding inconclusive evidence. In their study, they attempted to merge a variety of theoretical approaches with a group of

processes-related measures that prior studies did not employ, finding that these new contextual variables have indeed some meaningful influence on sentencing outcomes. Therefore, it would be too early to dismiss the influence of ecological variables on sentencing outcomes as trivial. So far, only a handful of contextual variables have been examined in the extant literature.

The first group of contextual variables includes court-organization level factors. The most often examined variable in the previous studies is jurisdictional size. The court community perspective states that the smaller the size of the court, the more severe the punishment, because the larger urban courts are more tolerant of social deviance and less visible from the outside world (Einstein et al., 1988). Prior studies found that there was indeed a rather consistently negative relationship between the size of the court community and the severity of sentences (Johnson, 2005; Kramer & Ulmer, 2002; Ulmer, 1997; Ulmer & Johnson, 2004), providing support for the argument put forward. Another closely related, but distinct, measure is urbanism-that is, the degree to which the specific jurisdiction is urbanized. Early sentencing research premised on the organizational perspective submits that sentencing outcomes in urban courts would be substantially less severe but also would be more likely to be contingent on extra-legal factors; this was not only because the administration of law was bureaucratic and rational (Flemming et al., 1992; Nardulli et al., 1988), but also because the urban contexts of the court community, characterized by low visibility and accountability, led to an informal local legal culture which further results in sentencing largely based on extra-legal factors (Einstein et al., 1988). The extant research, however, has produced mixed evidence, with one group of

studies reporting a significant effect of urbanization on sentencing (Myers & Talarico, 1986; Steffensmeier et al., 1993) and another body of research showing the opposite (Hagan, 1977).

Yet another organizational contextual variable that prior studies examined is caseload. The literature premised on the organizational perspective argued that court actors from a court community with heavy caseloads were more likely to spend less time on case processing and that as a result more lenient sentences would be imposed in such court communities (Dixon, 1995). Some prior studies found partial support for the inverse relationship between caseload and the odds of incarceration (Ulmer & Johnson, 2004; Ulmer et al., 2011b) or sentence lengths (Ward, Farrell, & Rousseau, 2009), while other studies uncovered no or a positive relationship between caseload and sentence severity (Kim, Cano, Kim & Spohn, forthcoming).

Researchers employing a contextual approach also have examined the impact of court environmental factors, including racial and ethnic composition and the level of socio-economic disadvantage in the community, on court outcomes. A review of the prior studies suggests that considerable attention has been devoted to the issue of to what extent the racial and ethnic composition of the court communities and, to a lesser degree, various types of political values in the community, shape sentencing outcomes and processes (Baumer & Martin, 2013). First, premised on the racial threat perspective, researchers investigated whether the proportion of racial minorities in the community (often conceptualized as the percent of African Americans) is associated with the level of

sentence severity. This group of studies yielded mixed findings, as some studies found the predicted positive relationship between the size of racial minority and sentence punitiveness (Bontrager, Bales, & Chiricos, 2005; Myers & Talarico, 1987; Ulmer & Johnson, 2004; Weidner, Frase, & Schultz, 2005), but other studies failed to find this relationship (Britt, 2000; Kautt, 2002; Ulmer & Kramer, 1996; Weidner & Frase, 2003). More recent empirical studies, however, attempted to examine more nuanced theses underlying the racial threat perspective. To illustrate, Wang and Mears (2010a) discovered that the relationship between percent Black and sentence severity is actually curve-linear, rather than linear, demonstrating that there is a tipping point where the proportion of racial minorities in the community starts to have a pronounced effect on judicial decision-making. In another study, Caravelis et al. (Caravelis, Chiricos, & Bales, 2011) also showed that a dynamic-, not static-measure of racial threat, represented as the increase in percent Black, explains the odds of being sentenced as habitual offenders in Florida (see also Wang & Mears, 2010b).

In a similar context, a body of research investigated whether the level of crime in a court environment was associated with punitiveness. Tyler and Weber (1982) made a case that the level of formal social control is a function of crime rates and fear of crime in specific locations, suggesting that crime rates should be positively correlated with overall punitiveness. This perspective, sometimes referred to as the crime threats perspective, is theoretically premised on the functional perspective, while the racial threat perspective discussed above is conceived as based on conflict perspective. This position is premised on the idea that tangible threat resulting from the increasing crime leads to fear of crime,

which in turn translates into a higher level of punishment imposed. Indeed, Myers and Talarico (1987) found that counties with higher crime rates were more likely to impose severe sentences, as compared to their respective counterparts. Jacobs and Helms (1996) also discovered that crime rates predicted the level of aggregated imprisonment rates, but the association was curve-linear (but see Britt, 2000). Baumer and Martin (2013) suggested that crime threat was mediated by the fear of crime, which was connected with violent crime rates, rather than general crime rates. Overall, there is rather consistent evidence in support of the predicted positive relationship between crime rates and the level of punitiveness in formal social control (see Weidner et al., 2004).

Local electoral politics is also frequently discussed in the court community perspective. It has been argued that between-court variations are wider in states or counties in which the judges or prosecutors are locally elected and in counties with more conservative political views (Kramer & Ulmer, 2004). Although several studies found null or equivocal empirical results for the effect of political conservatism on criminal sentencing (e.g., Baumer & Martin, 2013; Fearn, 2005; Jacobs & Carmichael, 2004; Johnson, 2005; Ulmer & Johnson, 2004), the results of a relatively larger body of studies suggested that local conservative politics are associated with more punitive sentencing outcomes in general (Helms & Jacobs, 2002; Huang et al., 1996; Johnson et al., 2008; Smith, 2004) and may also amplify racial and ethnic disparities in sentencing (Helms, 2009; Helms & Jacobs, 2002).

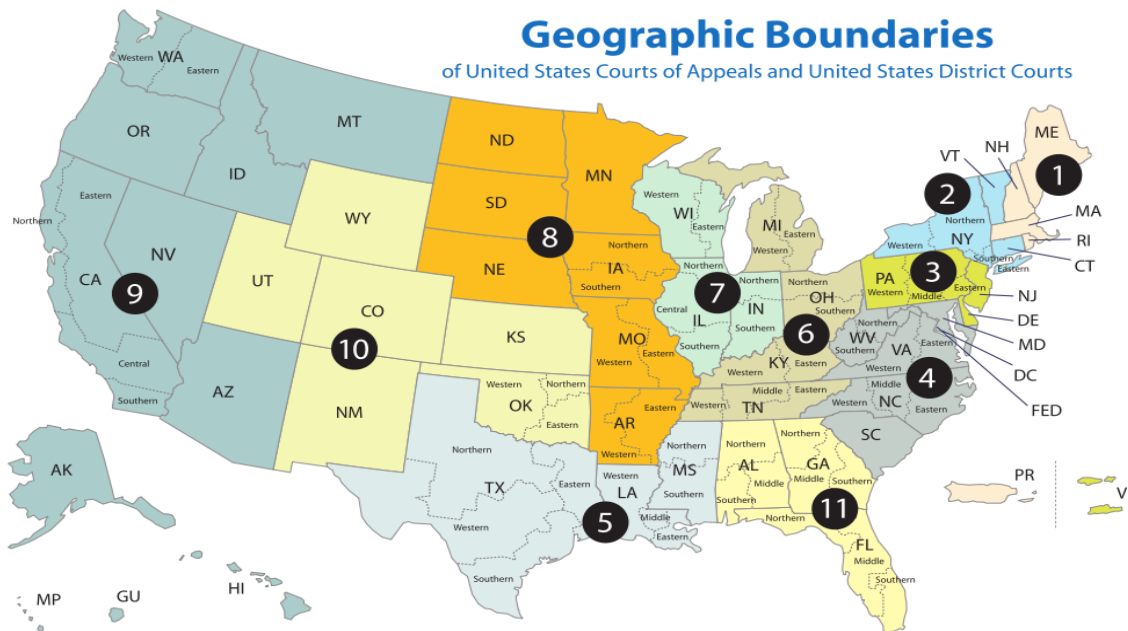
Socio-economic disadvantage (i.e. poverty and unemployment rates) is also one of the frequently tested theoretical constructs in social control literature and has a long tradition in its theoretical explanation (see, for example, Rusche & Kirchheimer, 1939). According to this economic threat perspective, economically disadvantaged groups are more likely to be perceived as threatening existing social orders, thus requiring more repressive responses from elites (Quinney, 1970; Spitzer, 1975). Recent sentencing literature also documents that offenders adjudicated in court environments with higher levels of poverty receive severe sentences, as compared to their counterparts adjudicated in jurisdictions with less poverty (Chiricos & Delone, 1992; Myers & Talarico, 1987). Conversely, there exists a body of works reporting a null relationship between socioeconomic conditions and sentence severity (Britt, 2000; Jacobs & Helms, 1996; Ulmer & Johnson, 2004).

The Case of Jurisdiction Variation in Federal District Courts

In order to understand between-court variation in the U.S. federal courts, it is imperative to have a good grasp on the unique organizational structure concerning the U.S. district courts. The federal jurisdiction consists of three layered courts: district courts at the bottom, circuit courts in the middle, and the U.S. Supreme Court at the top. Currently there are 94 District courts nationwide, including four district courts in the U.S. territories, with each state having one to four districts. These district courts again belong to 12 circuit courts, with the jurisdiction of a circuit court encompassing three or more states (See www.uscourts.gov for more information). The relationship between the district

court and the circuit court is complex; although most of the actions regarding sentencing occur at the district court, as there are not many cases which go before the circuit courts for appeal, the circuit courts also play an important role in many respects (see Rubin & Bartell, 1989). Most importantly, each circuit court is authorized to establish policies regarding its own appeals, thereby indirectly influencing the level of discretion of district judges. In other words, the discretion afforded to the district judges is supposed to vary, as there is a unique standard being applied to sentencing decisions made in each district (see Kautt, 2002). Secondly, each circuit court is also authorized to come up with its own calendar system, which affects how cases are processed in district courts (see Ulmer 1997 for a similar argument). Finally, the circuit courts also develop their own case law, thereby maintaining their unique ways of interpreting specific cases.

Figure-3 Federal Court Structures



Source: http://www.uscourts.gov/court_locator.aspx

Even though one might argue that the federal courts are not adequate units of analyses from the court community perspective, because of their bigger sizes and the selection methods of their court actors (see, for example, Johnson et al., 2008, p. 755), the federal jurisdiction is no less likely than the state courts to be contextually dependent, due to the ways in which each district and circuit court operates (Kautt, 2002). As suggested, it is true that federal judges are not elected by popular vote, which is unarguably one of the most important pathways upon which the effects of environmental factors exert their influence (Einstein et al., 1987). In federal courts, however, each circuit is authorized to implement its own policies and the district courts also have remarkable independence when it comes to forming policies. And the formation and implementation of policies are said to be influenced by broader sets of court environments (Ulmer, 1997). Consistent with these observations, Richardson and Vines (1970) claimed that “the location of federal courts throughout the states and regions renders them unusually susceptible to regional and local democratic forces” (p. 10). Providing support for this, research conducted prior to the implementation of sentencing guidelines demonstrated that there indeed existed significant between-district variation in sentencing outcomes and case processing strategies. In an early study, Sutton (1978) found that the predictors associated with sentence severity significantly differed across district and circuit courts. Einstein (1987) also discovered that federal prosecutors’ offices and their case processing strategies significantly varied across district courts. Reflecting this concern, it was claimed that these inter-jurisdictional disparities in sentencing and case processing were one of the reasons spurring the federal sentencing reforms (Breyer, 1988).

Prior studies on jurisdictional differences in federal courts subsequent to the implementation of federal guidelines painted quite a consistent picture, showing that federal districts still meaningfully varied in terms of their case processing strategies and sentencing severity. Earlier studies conducted in the 1990s mainly employed qualitative methods or rather simple quantitative analyses to describe different case processing strategies and sentencing practices based on only a small number of district courts. For example, Nagel and Schulhofer's (1992) study highlighted three district differences in case processing strategies following the implementation of the federal guidelines. The authors of this influential study found that there were systematic inter-district differences in charging and plea bargaining processes, including the application of substantial assistance departures (see also GAO, 1992; Maxfield & Kramer, 1998). The inter-district variability uncovered was, the authors concluded, attributable to different levels of commitment to the implementation of the federal guidelines and the lack of tight policies governing prosecutorial discretion. Two official reports published by the USSC also showed that there was a substantial amount of disparity between federal district courts in sentencing outcomes and the applications of mandatory minimums (USSC, 1991a, 1991b). In another early study, Weisselberg and Dunworth (1992) also demonstrated that the federal districts significantly vary in terms of their case processing strategies, such as plea rates and sentencing outcomes; thus, they concluded that the impact of guidelines was not homogenous across the federal district courts. Their conclusion is well encapsulated by the following quote: "A showing that different district and cases are

subject to different stresses is, in itself, significant, because it suggests that the guidelines mean different things in different contexts” (Weisselberg & Dunworth, 1993, p. 27).

More recent studies conducted since the year of 2000 tend to utilize data on all of the districts using more advanced statistical techniques, such as multi-level modeling. One of the first examples was Kautt (2002). Using the federal data from FY1999, she discovered that sentence lengths varied across both the district and the circuit courts, providing support for the core argument from the court community perspective. However, she failed to find any significant relationships between some of the important theoretical variables and sentencing outcomes; to be more precise, the results showed that there were no significant relationships between the proportions of Blacks, Hispanics, and unemployed in the district and sentence lengths. In addition, she also found no relationships among caseload pressure, the appeal rates of district courts, and sentence length. By contrast, she found that some aggregated measures of case processing outcomes (i.e. the rates of substantial assistance departure and guidelines-compliance) successfully predicted sentencing outcomes. She concluded that the significant inter-jurisdictional differences uncovered could be better accounted for by the explanatory factors derived from the social worlds perspective, even though the theoretical approach does not provide “a universally applicable explanation for the differential effects by jurisdiction” (Kautt, 2002, p. 659).

Using data from FY1997 to FY2000, Johnson and his colleagues (2008) examined whether district-level contextual factors could account for between-district

variations in the likelihood of departures. The results revealed that some contextual variables, for instance caseload pressure, influenced the odds of both the judge-initiated and the prosecutor-initiated downward departures. However, most contextual variables were differentially linked to the departures; judges from districts with lower levels of socioeconomic disadvantage and higher levels of political liberalism were more likely to grant the judge-initiated downward departures, but the same pattern was not observed in the application of substantial assistance departures. Furthermore, a defendant who was sentenced in a district court with a higher percentage of African Americans residents was less likely to be awarded the substantial assistance departure, but this relationship did not emerge for judge-initiated downward departures. Johnson and his colleagues also investigated whether contextual variables moderated the effects of extra-legal disparities in the applications of departures, finding that both racial/ethnic composition and, especially, the economic conditions of districts significantly reduced the odds of awarding downward departures to racial and ethnic minority defendants. Notably, these significant cross-level interactions were more evident with the judge-initiated downward departures. These findings led the authors to conclude that the federal courts differ in the propensity to depart due to complex social dynamics linked to the structural features of the federal court community.

Ward, Farrell, and Rousseau (2009) conducted a study investigating whether racial group balance in representation among courtroom actors in federal district courts is associated with sentencing disparity. In this unique study, they based their main argument on what is known as the power-balance perspective (Jackson, 1992) and measured the

Black index as the ratio of Black courtroom actors and Black residents, using the data from FY2001 to FY2002. They found that the prosecutor Black index accounted for variation both in incarceration and sentence length decisions; a district with a higher level of Black prosecutor representation was less likely to impose incarceration and more likely to mete out shorter sentences in general. They also uncovered interesting interactions between these district-level racial group representation variables and the defendant's race. Specifically, the results showed that Black judges and Black prosecutors significantly reduced the negative impact of being a Black defendant when deciding whether a defendant should receive a prison sentence. Overall, this study significantly expanded upon prior literature on race relationships and offered a promising avenue to contextual variation in sentencing outcomes in federal jurisdictions.

Wu and Spohn (2010) examined between-court variation in sentence lengths and the odds of downward departures among three selected districts: Minnesota, Nebraska, and Southern Iowa, using the data from FY1998 to 2000. Unlike other studies previously discussed, their study utilized the unique data set, which included detailed information on the defendants and interviews with the judges and prosecutors. The results revealed that two districts—Minnesota and Nebraska—were similar in terms of the average sentence lengths imposed and the source of disparity, such as substantial assistance departures, as compared to the district of Southern Iowa. They also found that there were both similarities and differences in the predictors of these outcomes. Their study provided rather mixed evidence regarding the uniformity of sentencing, but the weight of evidence,

according to the authors' conclusion, "clearly raise[s] questions about the validity of the assumption of uniformity in the federal sentencing process" (p. 316).

Ulmer, Einstein and Johnson (2010) focused their theoretical interests on what is better known as the trial penalty, examining whether the trial penalty significantly differs across federal districts. Based on data from FY2000 to FY2002, the authors found that there indeed existed substantial degrees of trial penalty—that is, the difference in sentencing outcomes between those convicted by trials and guilty pleas. More importantly, meaningful variation also existed in the trial penalty across the federal districts. In line with the authors' expectations, the trial penalty was more severe in districts with higher caseload pressure and lower trial rates. Premised on the findings, the authors concluded that the trial penalty was at least partly related to the organizational concerns regarding expedient case processing.

Finally, Feldmeyer and Ulmer (2011) used federal sentencing data from FY2000 to FY2002 to test the racial threat hypothesis. Even though some empirical attention had been paid to the issue of racial and, to a lesser degree, ethnic threats in sentencing research in state courts (Johnson, 2005; Myers & Talarico, 1987; Ulmer & Johnson, 2004), there were few studies conducted on this topic in the federal district courts (but see Johnson et al., 2008; Kautt, 2002). In particular, there was limited attention paid to what is called the "Hispanic threat" in federal district courts. The authors of the study concluded that racial threat theory did not apply to federal sentencing, as there was no relationship between the percent Black in the district and sentence length. However, the

proportion of Hispanic residents was related to sentence severity, but the direction of the relationship was opposite to that predicted by the racial threat theory; Hispanic defendants received more severe punishment in districts with smaller proportions of Hispanic residents.

To summarize, all of the studies reviewed so far point to the existence of inter-district variation in sentencing outcomes and processes, further demonstrating that the ecological contexts of federal court community may significantly affect judicial decision-making even under the restrictive sentencing schemes in the pursuit of uniformity in sentencing. One thing to note, however, is that the studies discussed above relied on federal data collected prior to the landmark decision in *U.S. v. Booker* (2005) and little is known about inter-district variation in sentencing practice following the *Booker* and *Gall* decisions. Given the documented wide variation in sentencing practice across the U.S. district courts in the pre-*Booker* era, some commentators expressed concern that the decisions would lead to increased discrepancies across districts; they called for a systematic inquiry looking into the between-court variation in sentencing outcomes in the wake of these major policy changes (Bowmen, 2006; Hofer, 2007).

There is some preliminary evidence confirming that the inter-district disparity in sentencing outcomes has indeed increased in the wake of *Booker* and *Gall*. (Bowmen, 2006; Farrell & Ward, 2011; Ulmer et al., 2011a). For example, Farrell and Ward (2011) sought to replicate their findings on the impact of racial group representation among courtroom actors on sentencing outcomes in the post-*Booker* period. Using the federal

data from FY2006 to FY2008 in an aggregated form, their study found that, with the exception of caseload pressure, none of the district-level covariates accounted for inter-district variation in sentence length. However, the study did reveal that some variables were associated with guidelines adherence rates, measured as the percent of within guidelines range sentences; districts with a higher level of political liberalism and higher violent crime rates were more likely to depart from the guidelines.

Finally, turning to the result of their direct interest—the effect of courtroom workgroup representation—it was shown that none of the variables explained between-district variation in sentence length and guidelines adherence rates. But, for Black defendants, judges were more likely to depart from the guidelines in a district with a higher level of African American prosecutors. In a similar context, Fischman and Schanzebach (2011) focused on examining the impacts of the standard of review and the aggregated judicial characteristics on sentencing outcomes. Using the federal data set from FY1991 to FY2007, they classified the whole 16-year period based on the level of judicial discretion allowed in accordance with the standard of appellate review, and investigated how the impact of aggregated judicial characteristics, such as political affiliation and pre-guidelines sentencing experience, differed across these different standards of review. The results suggested that the effect of aggregated judicial political affiliation was stronger under the more relaxed standard of review—that is, a deferential review, as compared to the more stringent review standard—a *de novo* standard. However, the effect of aggregated pre-guidelines sentencing experiences showed a modest direct

effect, not resulting in any significant interaction between the level of discretion and judicial characteristics.

Another study that examined changes in sentencing practices following Booker and Gall was conducted by Ulmer and his colleagues (2011a); they focused on extra-legal disparity and the inter-district variation in the effects of extra-legal disparity. Using the federal sentencing data set from FY2002 to FY2009 and a series of hierarchical regression models across four time periods (Pre-Protect, Protect, Post-Booker, Post-Gall), they found signs of relative stability in the effect of race/ethnicity and, more importantly, in the variation across district courts in the impact of extra-legal variables. Based on these findings, the authors concluded that the sentencing practices in federal district courts following Booker and Gall can be better characterized by stability than change. However, it should be noted that the methodology this study employed, including the inclusion of departure status as a control, was seriously challenged by some commentators (see Engen, 2011; Paternoster, 2011; Rehavi & Starr, 2012), who argued that the analysis would have yielded different results if the alternative methodology had been used. Meanwhile, more convincing evidence was provided by a study conducted by Kim and his colleagues (Kim et al., forthcoming). Based on the federal data from FY2000 to FY2010, the study demonstrated that the impacts of Booker and Gall had heterogeneous effects on sentence severity across federal district courts. More specifically, they found a reduction in sentence severity following Booker and Gall and a systematic pattern across federal districts in the degree to which sentence severity decreased: the reduction in sentence severity was smaller in districts characterized by higher proportions of Blacks in the

community and higher levels of economic disadvantage. Notably, political conservatism had an overall direct and positive effect on sentence severity, as some previous studies suggested, but the variable failed to moderate the impact of Booker and Gall in a significant way.

Summary of Prior Literature

This chapter reviewed three broad bodies of research in order to set the stage for addressing the specific questions that this dissertation focuses on. Beginning with a brief review on the backgrounds of the sentencing reform movement, the review suggested that the federal experiment was a part of bigger criminal justice reform movements during the 70's and 80's, even though the federal sentencing reform differed in many ways from other state experiences. What sets the federal guidelines apart from other state counterparts was the excessive rigidity which constrains judicial discretion and the extreme harshness in sentencing, which in turn created many controversies. Given these criticisms levied against the federal guidelines, a series of legal contingencies followed, reshaping sentencing policies and practices in the federal district courts. Notably, the most important landmark decision was the Booker decisions, which rendered the federal sentencing guidelines effectively advisory. The review of the post-Booker literature suggested that the field is still struggling with teasing out the key implications of the decisions.

The review of the second group of studies focused on the historical, theoretical, and empirical literature on gender and sentencing. First and foremost, the extant literature

clearly demonstrated that women defendants were treated with greater leniency in sentencing over their male counterparts, even though the conditioning effects involving gender turned out to be not as consistent as the main effect of gender. One important limitation that stood out from the review concerns that previous studies exclusively focused on examining the interaction effects between gender and the immediate circumstances surrounding gender (i.e. crime type and a variety of family situations), neglecting broader socio-political environments that directly and indirectly influence judicial decision-making. Therefore our knowledge on the potential relationship between gender and the socio-political environments of the court communities is limited. Meanwhile, the review of the theoretical literature on gender and sentencing underscored the importance of gender norms or gender stereotypes in accounting for the female discount that the previous studies found. However, the extant bodies of work appeared to pay relatively scant attention to the nature of the gender norms and the process by which these gender stereotypes impact judicial decision-making.

Finally, the third group of literatures discussed deals with the topic of social ecology and sentencing. In this review, the dominant theoretical perspective, the court community perspective, suggested the possibilities that courts vary both in terms of sentencing outcomes imposed and case processing strategies and that the implementation of formal sentencing policies is always filtered through the substantive rationality built into the local legal culture within the court communities. The review of the previous empirical studies on social ecology and sentencing provided strong support for the former prediction, demonstrating the existence of between court-variations. However, there

appears to be a clear gap in the literature as to the latter prediction on whether and how the changes in formal sentencing policies have differential impacts on sentencing practices and case processing strategies across different court communities. Turning to the review of the studies conducted in the federal district courts, the literature review clearly indicated that federal district courts vary substantially in term of their adherence to guidelines and the level of punishment imposed. It was further shown that the recent transition of the federal guidelines into the advisory guidelines seemed to increase inter-district variation in sentencing but it is fair to say that not much is known about the nature of this inter-district variability in the wake of Booker and Gall.

Taken as a whole, this dissertation attempts to address the gaps in the previous studies discussed so far by merging the literature on gender and sentencing with the literature on social ecology and sentencing. From a theoretical point of view, special attention will be paid to the theoretical foundation of the focal concerns perspective, symbolic interactionism, the view that the elements of the focal concerns perspective are “socially constructed” (Daly, 1994, p. 169) and arguments that “the causal mechanisms of sentencing decisions ... lie in the interpretive processes” (Ulmer, 2012, p. 8). That is, not only the concept of blameworthiness, but also the practical constraints and social costs associated with punishment toward women are subject to varying interpretations by the key players of the courtroom workgroup (Ulmer & Kramer, 2006), which is inevitably shaped by the ecological contexts of the court community (Eisenstein et al., 1988). Here in this dissertation, I will specifically focus on the two most important socio-political factors surrounding court communities, such as religious and political

conservatism. From an empirical point of view, this dissertation will also utilize a unique opportunity which occurred in the federal districts in the wake of Booker/Gall—the transition of the federal guidelines into the advisory guidelines. In doing so, two particular questions that the dissertation attempts to answer are whether gender disparity would increase or decrease following Booker/Gall and how the ecological contexts of the court communities further condition the potential changes in gender disparity.

CHAPTER 3: CURRENT FOCUS

In this chapter, I will lay out in detail the specific contexts of the two studies, which are the focus of my dissertation.

STUDY-1: Gender Disparity and The Impact of Ecological Contexts: Implications on Racial and Ethnic Heterogeneity

In the first part of the dissertation, I will examine the effects of two particular dimensions of the court community—that is, their religious and political contexts—on gender disparity, as they are closely linked to both gender stereotypes and punishment severity (Baumer & Martin, 2013; Bohm, 1991; King, 2008; Ulmer et al., 2008), and may alter gender disparity in sentencing (Helms & Jacobs, 2002). Some recent scholarship emphasizes the role of religious and political conservatism in shaping courtroom decision making (Garland, 2001; Jacobs & Carmichael, 2004; King, 2008; Savelsberg, 1994). Even though they are closely related, religious and political conservatism are “independently predictive of state social control” (King, 2008: 1354) and may have differential effects on associated beliefs and actions of court actors (Baumer & Martin, 2013; Gross, Medvetz, & Russell, 2011; Thorne, 1990). Therefore, here I will focus on the moderating effects of both religious and political contexts of the court community on gender disparity. Specifically, I will examine whether the impact of gender varies depending upon the level of religious or political conservatism in the community in which the court is located. In doing so, particular attention will be devoted to the

possibility that the conditioning effect of the ecological contexts of the court community on gender is further shaped by the race and ethnicity of defendants.

Religious Context and Gender Disparity

Even though the role of religiosity or the institution of religion on the level of social control is well established (see Durkheim, 1985), the relationship between religious contexts and criminal sentencing in specific remains one of the most under-researched topics in ecology and sentencing literature (Ulmer, 2012). The study by Ulmer and his colleagues, in which they tested a set of hypotheses involving the direct and conditioning impacts of religious contexts of court community on sentence severity, is the only exception to date (Ulmer, Bader, & Gault, 2008),

Two broad roles of religious contexts of court community could be derived from Ulmer et al.'s study: the first hypothesis suggests that specific types of religion prevalent in the court community are important in shaping court outcomes. For instance, Garland (1990) claimed that fundamentalists may be more likely to be punitive toward offenders, because they tend to interpret the Bible literally, thus placing more emphasis on individual responsibility for offenders' misbehaviors (see Luper, Hopkinson, & Kelly, 1988). A competing stance for the role of religious contexts would be the argument that the level of religious homogeneity, not the specific types of religion, is more important in exercising social control in society. This position is strongly embedded within the argument made by Durkheim, in which religious consensus, represented by religious homogeneity, was the one which set the stage for strong formal social control (see Stark,

Doyle, & Kent, 1980). In the study, Ulmer et al. (2008) found that counties with higher levels of religious homogeneity were more likely to incarcerate defendants. In addition, they also found that both the religious heterogeneity and the percent conservative Protestant intensified the effects of offense severity on the odds of incarceration. Finally, counties with a higher Christian heterogeneity were especially more likely to be punitive toward defendants with higher criminal history scores. In sum, this study demonstrated that the religious contexts influenced court actors' perceptions and shaped court outcomes in a complex way.

In the present study, I will focus on investigating the conditioning effect of the specific type of religion—Conservative Protestantism—rather than religious heterogeneity, because Conservative Protestantism may be most closely related to gender stereotypes and patriarchal culture in the community, which supposedly is related to gender disparity in sentencing. As noted earlier, the dominant theoretical perspective in courtroom decision-making—the focal concerns theory—placed an almost exclusive emphasis on societal stereotypes in explaining extra-legal disparity. For that reason, it makes sense to assume that a specific type of religious context, not Christian homogeneity, more closely fits with gender stereotypes. In that regard, prior research in religious studies, feminist literature, and criminology described Conservative Protestantism as a patriarchal religion (Bartkowski & Read, 2003; Vieraitis, Britto & Kovandzic, 2007) or a conservative religion (Fearn, 2005; Ulmer et al., 2008). More specifically, the tradition of Evangelical Protestantism emphasizes a wife's submission to her husband and depicts women as subservient to men (Blanchard & Prewitt, 1993). Further, prior research has revealed that

traditional gender roles are prevalent within conservative Protestantism in that females are not expected to be involved in the labor force, and even in the household, there is a very clear gender division with regard to the tasks for husbands and wives to perform (Bartkowski, 1997; Peek et al., 1991).

The specific direction of the impact of religious context on gender disparity, however, may be in both ways. The first scenario would be the possibility that female offenders in court communities characterized by higher levels of religious conservatism are in fact subject to harsher sentences, as compared to women defendants from less conservative religious contexts, thus leading to a relatively smaller gender disparity in sentence severity. This is because such court environments may be more in favor of traditional paternalistic views concerning appropriate roles for women. Thus, the misbehaviors of female defendants, especially married women with dependents, may be considered more blameworthy, as compared to those of female defendants in court communities with a less conservative mood. Taken together, this position places a heavy emphasis on how court actors in a religiously conservative community would view female defendants negatively, suggesting that the element of blameworthiness mainly drives judicial decision-making (see also Baumer & Martin, 2013, p. 171).

The alternative scenario would be that female defendants adjudicated in communities with higher levels of religious conservatism may be treated more leniently than their counterparts from less conservative communities, and, thus that there would be a larger gender disparity in these court settings. What complicates dynamics involving

the treatment of female defendants is the fact that judges in more religious conservative communities may also encounter strong pressure from the communities to protect families and children, and to value the role of husbands as a form of informal social control. In fact, this expectation is more in line with the theoretical argument made by Kruttschnitt (1982, 1984) and Kruttschnitt and Green (1984), in which female defendants were argued to be accorded leniency in formal social control because of a higher level of informal social control in their lives. If this were the case, gender disparity favoring woman defendants may be amplified by religious conservatism such that the extent of gender disparity in sentencing is larger in jurisdictions with a higher level of religious conservatism. To summarize, this position would be contrasted with the former hypothesis in that the element of practical concerns, rather than blameworthiness, is the primary influence on judicial decision making.⁵

Political Context and Gender Disparity

The relationship between political contexts and social control has been well discussed in theoretical literature (see Chambliss, 1964; Foucault, 1977; Savelsberg, 1994). It has long been argued that social control is inseparable from the institution of politics and politics are also deeply embedded within other social arrangements such as class or racial divisions in society (Garland, 2001). More recent literature also

⁵ Another possibility would be that the impacts of the first two elements of focal concerns may cancel them each other, thus not producing any inter-district discrepancies in gender difference. That is, a court community with a high level of religious conservatism may be more likely to view the female offenders as more blameworthy, but at the same time, to care about the practical constraints resulting from punishing these female offenders.

emphasized the importance of the role that politics has played in the get-tough policies observed since the late 1970s. Beckett (1997) made a strong case that a law-and-order rhetoric had been utilized by conservative politicians over the last three decades that viewed the lack of individual moral responsibility as the main cause of criminal behavior, thus calling for punitive responses to those amoral behaviors. Garland (2001) also claimed that a recent conservative response to criminal behavior in the U.S. was largely shaped by conservative politicians who sought to mobilize the cultural sensibility of crime which characterizes a modern society.

Reflecting these lines of arguments, prior studies found that political forces at the national (Chambliss, 1994), state (Jacobs & Carmichael, 2004), and local (Eisenstein et al., 1989) levels had distinct impacts on the level of formal social control (Jacobs & Helms, 1996). As one of the most extensive studies conducted on this topic, Jacobs and Helms (1996) discovered that political contexts indeed had a strong effect on aggregated incarceration rates. More specifically, they found that incarceration rates had increased in accordance with the year of the presidential election since 1965, thus providing support for the election cycle hypothesis. At the same time, they further found that Republican candidates had a bigger impact on the level of incarceration, net of the election cycle, thus offering evidence for the partisan argument.

In the sentencing literature more attention has been directed at the role of local politics in judicial decision-making (see Eisenstein et al., 1988; Levin, 1972). As a specific aspect of formal social control, the criminal sentencing process is also

considered fundamentally political; not only is political culture a central aspect of local legal culture (Eisenstein et al., 1988), but local politics play an influential role in selecting courtroom actors (Eisenstein & Jacob, 1977; Ulmer et al., 2008). Some earlier studies in this tradition focused on the impact of judges' political affiliation on sentencing outcomes (see Myers & Talarico, 1987), but more recent studies emphasized the impact of the political contexts of court communities on judicial decision-making (see Ulmer, 2012). Although several studies found null or equivocal empirical supports for the effect of political conservatism on criminal sentencing (Fearn, 2005; Jacobs & Carmichael, 2004; Ulmer & Johnson, 2004), others found that local conservative politics were associated with more punitive sentencing outcomes in general (Baumer & Martin, 2013; Huang et al., 1996; Johnson et al., 2008) and also amplified racial and ethnic disparities in sentencing (Helms, 2009; Helms & Jacobs, 2002). Therefore, the political contexts of the court community have been said to be closely linked to punitiveness in response to crimes (Jacobs & Carmichael, 2001; Jacobs & Helms, 1996).

Here I will lay out two equally plausible, but opposing, situations as to how local political conservatism might condition the extent of gender disparity in sentencing. First, contextual factors concerning political conservatism may reduce gender-based sentencing disparity. Conservatives use law-and-order appeals to attract voters, and those who are politically conservative are more likely to endorse the law-and-order ideology (Beckett, 1997). One of the traditional strategies that Republican politicians employ to expand their political appeal to the working middle-class has been to campaign on a law-and-order agenda (Chambliss, 1994). And, the politics of law and order put forward a number

of gender-blind sentencing policies, such as mandatory minimums for drug offenses, which, according to some legal scholars, significantly contributed to a decreased gender gap in sentencing outcomes, especially in federal district courts (Chesney-Lind, 2002; Mauer, 2013). Indeed, recent research found that the degree to which mandatory minimums are implemented as intended depends on the political contexts of court communities (Ulmer, Kurlychek, & Kramer, 2007). Thus, the degree to which the law-and-order movement is prevalent in court environments may undermine the traditional chivalry hypothesis by stressing individual responsibility and equal justice, which may ultimately result in a smaller gender gap in punishment. In other words, gender disparity may be significantly reduced in the mostly politically conservative court environments.

However, it is also equally plausible that gender differentials in sentencing outcomes may be larger, with females receiving more lenient sentences compared to their male counterparts in court communities characterized by a higher level of political conservatism. This would be due to the fact that political conservatives typically hold more traditional paternalistic views as conservative political climates reinforce the traditional gender roles by not endorsing policies in support of gender equality (Vieraitis et al., 2007). For the similar reason, conservative political moods are more likely to define women, not men, as the primary caretaker and put more emphasis on the caregiving role than breadwinning role in maintaining family (Daly, 1987a; 1987b). These overall sentiments may manifest themselves in chivalrous attitudes among criminal justice actors in formal social control settings, as the paternalistic culture would require informal social control imposed by husband and family. Thus gender disparity favoring

woman defendants may be amplified such that the extent of gender disparity in sentencing is larger in jurisdictions with a higher level of political conservatism and smaller in those with a lower level of political conservatism. Indeed, Helms and Jacobs (2002) partially tested this possibility and found that female defendants were accorded more preferential treatment over their male counterparts in courts with a higher level of political conservatism.

Gender Disparity and Ecological Contexts along Racial/Ethnic Lines

The theoretical underpinnings presented above assume that the religious and political contexts of court environments influence gender disparity, regardless of defendants' race and ethnicity. Yet, a more careful consideration of race-specific gender stereotypes (Krivo, Peterson, & Hagan, 2006) and different family circumstances across racial and ethnic groups (Wilson, 1987) may challenge this assumption for at least two reasons. Firstly, gender stereotypes may have different implications for racial/ethnic minority females than for white females. It has been documented that minority females, compared to their white counterparts, are stereotyped more negatively (Healey, 1997; Madriz, 1997; Portillos, 1998; Young, 1986); women of color are oftentimes described as dirty, hostile, superstitious, and independent (Brennan, 2006); they also are characterized as hyper-sexed and as welfare queens (Alexander, 2012). Relatedly, women of color, particularly Black females, are often depicted as unfit mothers (Huckerby, 2003), although they are more likely to be situated in the position of being the sole head of a household with child rearing responsibilities (Cherlin, 1992; Wilson, 1987). In sum, the

traditional perception of femininity, the core concept behind the chivalry hypothesis, is more likely to fit into gender roles for white females, and thus, compared to white female defendants, non-white female defendants may face harsher treatment from the criminal justice system (Klein, 1995).

Secondly, the punitive crime control policies, symbolized by the war on drugs, have mainly targeted communities of color (Alexander, 2012), and sentencing policies brought about by the war on drugs have resulted in a dramatic growth in incarceration of minority females especially for drug offenses (Chesney-Lind, 1995; Mauer, 2011; Tonry, 1996). For instance, there is a small group of studies which demonstrate that the drug sting operation is more rigorously enforced in an area where racial/ethnic minority offenders reside or hang around (Alexander, 2012; see Belenko & Spohn, 2014). Similarly, commentators noted that mandatory minimums are more likely to be directed toward offense types committed by racial and ethnic minorities (Becket & Sasson, 2000; Kautt & Spohn, 2002; Tonry, 1987). Indeed, researchers uncovered that mandatory minimums were more likely to be applied in court communities characterized by a higher level of racial/ethnic minorities (Ulmer, Kurlychek, & Kramer, 2007) and that mandatory minimums were more likely to be applied to racial and ethnic minority offenders (Caravelis et al., 2011). Considering this body of evidence, then, racial and ethnic minority defendants, regardless of their gender, may be subject to distinct social control climates, compared to their white counterparts.

All these theoretical arguments speak to the possibility that the interactions between gender and the ecological contexts may be further conditioned by the race and ethnicity of female defendants. For example, because Black females may be more likely to be viewed in violation of traditional gender stereotypes, they may be subject to harsher treatment as compared to white female offenders (Bickle & Peterson, 1991; Kruttschnitt, 1982), and this pattern may be more pronounced in a court environment where conservative religion and politics are more prevalent. On the contrary, it is also possible that the consequence of imprisoning the sole head of a household would have great consequences for minority women and thus judges may mete out less severe punishment to non-white female offenders (Daly, 1989; 1994). Likewise, this pattern may be more pronounced in a district court with a higher level of political liberalism, because court communities characterized by higher levels of political liberalism may be more concerned about the collateral consequences of incarceration of this kind. In that sense, I suspect that the inconsistent findings on the issue of whether chivalry bypasses women of color may be partly due to the lack of attention to the role of ecological contexts of court community in relation to gender and race.

Hypotheses

Drawing on the discussions above, five hypotheses are developed, focusing on gender disparity in sentencing. The first hypothesis addresses the main effect of gender; it predicts that female defendants will receive more lenient sentences than male defendants. The next two hypotheses focus on the contingent effects of local religious and political

conservatism on gender disparity in sentencing. As noted, there exist equally plausible, but opposing, possibilities regarding the directions of the potential interactions between gender and political/religious conservatism. Based on the empirical observation that the recent incarceration rates for female offenders has been on the rise (Mauer, 2013), in this dissertation, I expect that the traditional gender discount in sentencing outcome to be smaller in jurisdictions with higher levels of political or religious conservatism.

Hypothesis four concerns the possibility raised by Ulmer (2012) that religious and political contexts may interact to moderate gender disparity. The last hypothesis investigates if the contingent effect of religious and political conservatism on gender-based sentencing disparity is further conditioned by race and ethnicity of the defendants.

Given the race/ethnicity specific gender stereotypes, I expect that non-white female offenders adjudicated in court communities with higher levels of political/religious conservatism would enjoy smaller gender discounts relative to other non-white female offenders adjudicated in court communities with lower levels of political or religious conservatism. This expectation is also grounded on the argument that the main determinant of increasing female incarceration rates is at least partially related to the politics of the war on drug (Mauer, 2013) and communities of color are oftentimes the targets of law enforcement operations. Specific hypotheses are stated below:

Hypothesis 1: Female defendants will receive shorter sentences than their male counterpart.

Hypothesis 2: The level of religious conservatism in a jurisdiction will affect gender disparity such that female defendants adjudicated in jurisdictions with higher levels of religious conservatism will receive smaller sentence discounts relative to other female defendants adjudicated in jurisdictions with lower levels of religious conservatism.

Hypothesis 3: The level of political conservatism in a jurisdiction will affect gender disparity such that female defendants adjudicated in jurisdictions with higher levels of political conservatism will receive smaller sentence discounts relative to other female defendants adjudicated in jurisdictions with lower levels of political conservatism.

Hypothesis 4: Gender disparity will be least pronounced in jurisdictions with higher levels of both religious and political conservatism.

Hypothesis 5: The interaction effect between female and religious conservatism—the degree to which religious conservatism reduces the female discount—will be significantly greater for non-white female offenders, as compared to white female offenders.

Hypothesis 6: The interaction effect between female and political conservatism—the degree to which political conservatism reduces the female discount—will be significantly greater for non-white female offenders, as compared to white female offenders.

STUDY-2: Gender Disparity, Policy Changes, and Ecological Contexts of Court

Community

As noted earlier, the extant body of research on sentencing has demonstrated that some level of disparity in sentencing remains even in jurisdictions with sentencing guidelines (see Spohn, 2000; Wooldredge, 2009). However, what has been particularly lacking is an effort to evaluate the impact of the implementation of the reform movement by making comparisons between pre- and post-guidelines sentencing practices (Kramer & Ulmer, 2009; Spohn, 2009). Engen (2009) criticized extant sentencing research in this regard, maintaining that scholars should conduct research on the impact of repeal or relaxation of sentencing guidelines on sentencing practice in order to have a better understanding of the relationship between policy changes and sentencing practices and, in particular, to determine whether policy changes affect levels of disparity based on offender characteristics such as race, ethnicity and gender. With respect to gender, many sentencing guidelines prohibit considering both gender and gender-related factors that are closely related to the unique circumstances of females (i.e. pregnancy, child-rearing responsibility etc.), and thus there is reason to believe that the preferential treatment toward female defendants may have declined after the implementation of sentencing guidelines (Rader, 1993). Reflecting this concern, some commentators suggested that the implementation of the sentencing reforms may have disadvantaged women more than men (Chesney-Lind & Pollack, 1995; Daly, 1994; Nagel & Johnson, 1994).

Although there is some preliminary evidence in support of this assertion (see Mauer et al., 1999), the evidence is not conclusive due to the lack of research on this topic. In the current study, I will move the focus of the discussion to the influential policy change that recently occurred in the federal jurisdiction—the transition of the federal sentencing guidelines into advisory guidelines—, and investigate whether the recent shift has resulted in any changes in gender disparity. In doing so, special attention will be devoted to the main topic of this dissertation, examining the possibility that the magnitudes of potential changes in gender disparity are shaped by the ecological contexts of federal district courts.

Gender Disparity and Changes in Sentencing Policies

As suggested previously, there exist conflicting views on to what extent the sentencing reform movement could shape sentencing practice, with one perspective indicating a full-fledged change and the other suggesting a limited or more complex impact (see Kramer & Ulmer, 2009; Savelsberg, 1992). With regard to the current topic of interest, there are only a handful of empirical studies offering clues as to whether gender disparity varies following changes in sentencing policies. To begin, a group of relevant studies were conducted in Minnesota, where gender and other gender-related factors are prohibited in judicial decision-making: first, Knapp (1984) argued that overall sentencing disparities decreased in the wake of implementation of the Minnesota sentencing guidelines. More importantly, it was claimed that gender disparity also diminished, but the reduction was mainly achieved through increasing the severity of

punishment toward female defendants. Using a more advanced statistical technique with rich control variables, Miethe and Moore (1985) painted a little different picture, finding no gender disparity in either the pre- or post-guidelines eras in the imprisonment and sentence length decisions. However, they uncovered a nuanced change in the effect of gender following the introduction of the guidelines, finding that the effect of gender on sentence length was indirect through charge bargaining, which indicates less preferential treatment toward women in the post-guidelines era.

Finally, as one of the most influential studies on this topic, Koons-Witt (2002) found that gender alone did not have a significant impact on the likelihood of imprisonment either in the pre- or post-guidelines eras. However, women with dependent children were less likely to be sentenced to prison during the pre-guidelines era, though the effect dissipated in the years subsequent to implementation of the guidelines. But the interaction effect favoring female defendants with children reemerged in the later part of the guidelines implementation, when the guidelines were assumed to be effectively in place. Taken together, these four studies seem to suggest that the gender disparity in Minnesota during the pre-guidelines era was not pronounced but rather was nuanced in nature, and that the impact of the guidelines on gender disparity may have been minimal, especially in the long term.

Another important study was conducted in Ohio, where the judges retain wider discretion than in Minnesota. Using data from Ohio collected before and after the passage of the guidelines, Griffin and Wooldredge (2006) attempted to replicate the Koons-Witt

(2002) study: they found that female offenders were treated more leniently in the imprisonment decision, but not in the sentence length decision during both the pre- and post-sentencing guidelines eras. In addition, the interaction effect involving women with dependent children did not have a significant effect in the pre-guidelines era. But counter-intuitively, the same effect turned out to be significant such that women with dependent children were more likely to receive longer sentences in the post-guidelines era. Overall, the authors summarized their findings that extra-legal disparities changed little across the sentencing reform movement and this limited impact of the guidelines resulted from the wide discretion currently allowed to judges and the way in which the average sentence severity was constructed in Ohio.

Finally, Blackwell, Holleran, and Finn (2008) utilized a unique opportunity to examine whether changes in sentencing policies affect gender disparity in Pennsylvania where the judges retain a substantial amount of discretion and are not prohibited from taking gender into consideration. Starting off with the historical backgrounds for the Pennsylvania Sentencing Guidelines, the guidelines were initially instituted in 1982 and the Pennsylvania Supreme Court suspended the guidelines in 1987, which became effective a year later in 1988 after the Sentencing Commission quickly addressed the constitutional issues. The authors paid attention to these changes in guidelines implementation (i.e., the pre-suspension, suspension, and post-suspension periods), and examined how gender disparity changed across these three distinct periods. The results revealed that the effect of gender on the in/out and sentence length decisions was constant across the three stages; female defendants were less likely to be incarcerated and more

likely to have shorter sentences imposed as compared to their male counterparts. The authors concluded that the guidelines had little impact on gender disparity in sentencing outcomes in Pennsylvania.

All things considered, the review of the prior studies clearly indicates a limited impact of sentencing guidelines on levels of gender disparity. However, one thing to take note of is the fact that the sentencing guidelines that the studies are based on differ substantially in term of judicial discretion afforded to the judges and the extent to which gender and gender-related factors can be taken into account. In this regard, Koons-Witt (2000) called for another study focusing on the sex-based disparity from different guidelines schemes (see also Blackwell et al., 2008). More importantly, all of the previous studies failed to recognize an important insight that the impact of gender may not be invariant across different court communities even within a jurisdiction, even though the impact of the sentencing reform is argued to depend on the degree to which formal rationality is at odds with substantive rationality inculcated in the court communities (Savelsberg, 1992). Therefore, in what follows, I will move the focus of discussion to the potential changes in gender disparity in federal district courts, and, in doing so, specific attention will be devoted to the impact of the ecological contexts of federal court community on the changes in gender disparity.

Gender Disparity, Booker and Its Social Contexts: The Case of Federal Sentencing

One of the central goals of the federal sentencing guidelines is the reduction of unwarranted sentencing disparity (Kramer & Ulmer, 2009; Spohn, 2008). Reflecting this

goal, the SRA specifies gender as one of the factors that judges are precluded from taking into account in their sentencing decisions (28 U.S.C. §994(d)). Despite the intention of architects of the guidelines, surprisingly, it was suggested that gender disparity increased following the introduction of the guidelines. As the USSC put, “the gap in average prison terms between male and female offenders has widened in the guidelines era” (USSC, 2004, p. 127). Furthermore, the report noted that the magnitude of disparity was so large that sentence lengths for men were typically 25 to 30 percent longer than for women (USSC, 2004). Despite the unexpected strong gender effect, questions still remain over whether and why the gap increased rather than decreased, as the USSC report was not predicated on the pre-and post-comparison following the implementation of the federal guidelines.

Given the fact that data from the pre-guidelines era are not available, it is impossible to replicate the approach taken by Koons-Witt (2002) and Griffin and Wooldredge (2006) in the federal courts. Instead, the focus of the present study is directed toward investigating the impact of federal guidelines’ transition into advisory guidelines from presumptive guidelines and the effect of this change on gender disparity. To date, there is no specific study that examines this interesting possibility, even though a small body of research offers some hints on the potential relationship. To begin, Ulmer et al.’s (2011a; 2011b) analyses implied that the effect of gender did not change substantially in the post-Booker and post-Gall/Kimbrough periods, as compared to the pre-Booker period. Similarly, the USSC’s Booker report (2006) also suggested that the effect of gender had changed little. Two follow-up studies conducted by the Commission

(2010; 2012) revealed that there seemed to be a slight reduction in gender disparity following Booker, but the disparity reverted back to normal or increased following Gall, especially the later period of Gall (USSC, 2012). These studies suggest that changing the federal guidelines from presumptive to advisory status did not significantly or consistently influence how female defendants are treated in federal courts.

However, the fact that there are a number of limitations to the research conducted to date suggests that this temporary conclusion regarding the impact of Booker/Gall on gender disparity in federal district courts may be premature. First, these studies suffer from a myriad of methodological issues, thus calling into question the validity of their findings. For instance, the offender's departure status was employed as a control variable in each of all the studies reviewed, thus eliminating a substantial amount of variation in which the impact of Booker/Gall could occur. Given a recent methodological debate (Ulmer et al., 2011b v. Starr & Rehavi, 2013), a different model specification without the departure control may have yielded different findings regarding gender. In addition, these studies failed to control for any temporal variation which might have been present in specific time periods, even though these periods were relatively long and characterized by changes in social factors such as crime rates (see Levitt, 2004). Meanwhile, at the conceptual level, it also needs to be recognized that the main implication of the Booker and Gall decisions—a decrease in overall sentence severity (Fischman & Schanzenbach, 2012; Kim et al., forthcoming) —was not well appreciated in the previous studies. That is, any Booker/Gall research should start from the observation that the average sentencing severity has decreased and accordingly any changes in extra-legal disparities (i.e., those

based on race, ethnicity, and gender) need to be understood against this backdrop. In order to answer the question at hand in a methodologically sound way, a new framework upon which to investigate any potential changes in gender disparity following Booker and Gall is required.

What also is lacking in the research conducted to date on the implications of Booker and Gall in federal sentencing is an effort to measure the impact of Booker and Gall/Kimbrough in terms of ecological contexts. As suggested by the court-community perspective, any changes in sentencing policies are almost always filtered through informal sanction norms shaped by ecological contexts of court community (Eisenstein et al., 1988; Myers & Talarico, 1987). Following this logic then, it is quite plausible that the impact of Booker and Gall may have been distinct across different federal district courts. To date, there is very little research addressing the issue of inter-district variations in the effects of Booker and Gall. Although few, existing studies demonstrated the possibility that there were indeed some inter-district variations in sentencing outcomes following Booker and Gall (see Farrell & Ward, 2011; Ulmer et al., 2011b). One study formally tested whether the impacts of the Booker and Gall decisions were conditioned by the ecological contexts of the court community. In their study, Kim and his colleagues (forthcoming) discovered that the transition resulted in the decrease in sentence severity and had differential effects along the ecological contexts of court community. In particular, they found that the effect of the percent Black and disadvantage conditioned the effect of Booker and Gall, showing that the reduction in sentence severity was significantly smaller in districts with higher levels of African Americans and socio-

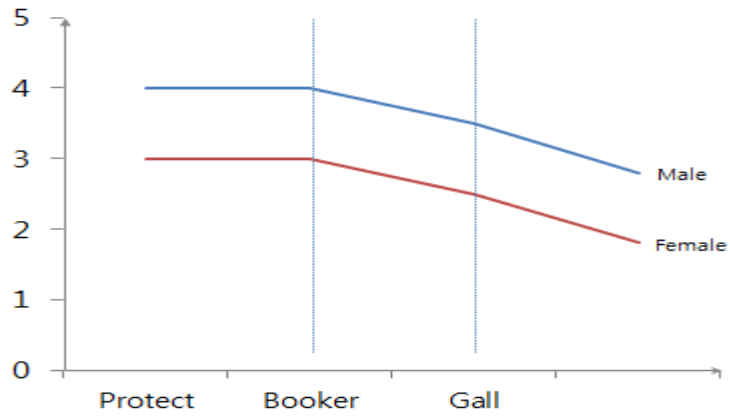
economic disadvantage. Combining the two lines of research on changes in gender disparity in the wake of Booker/Gall and the ecological contexts of federal district courts in shaping the change, it is possible to expect that any changes in gender disparity as a result of these Supreme Court decisions may vary significantly across federal district courts.

HYPOTHESES

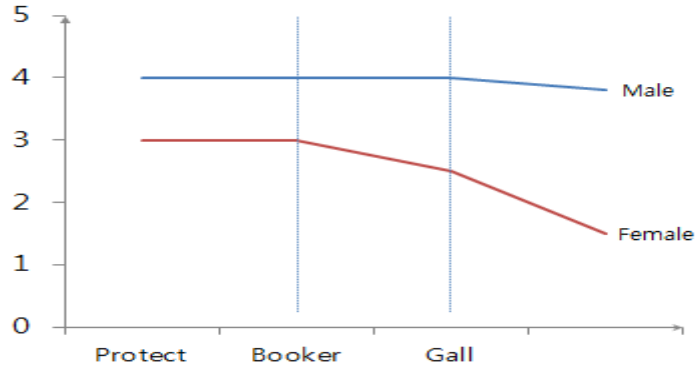
Building on the previous findings that there has been a reduction in overall sentence severity following Booker/Gall (Fischman & Schanzenbach, 2012; Kim et al., forthcoming), three potential situations could be inferred involving the patterns of gender disparity across Booker/Gall, as illustrated by figure-4 below.⁶ The first model suggests that there is no change in gender disparity in sentence severity following the Booker and Gall decisions. In this conceptual model, the reduction in overall sentence severity in the wake of Booker and Gall would be observed for both male and female defendants with approximately the same amounts. The second model expects that the decrease in overall sentence severity following Booker and Gall would result mainly from the reduction in sentence severity for female defendants. In this regard, Raeder (2006) makes a case that the Booker decision would allow the judges in the federal courts more flexibility to depart from the guidelines based on gender-related factors, which have more direct relevance to female defendants, and the judges are expected to utilize this opportunity.

⁶ It needs to be noted that the graphical representations offered in Figure-4 is only for the explanation purpose and not necessarily related to the analytic schemes for the current study.

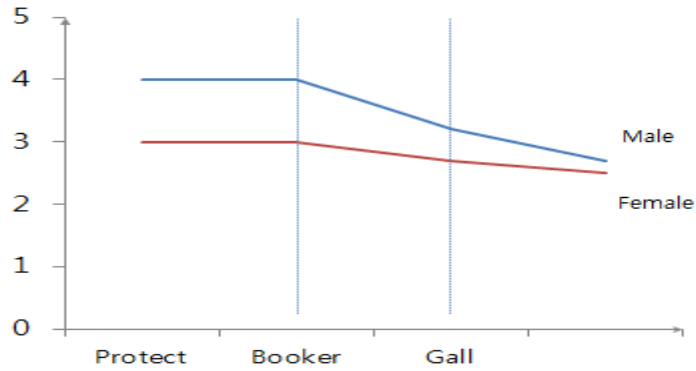
Figure-4 Conceptual Framework related to Hypothesis-1 and -2



<Model-1: no differences>



<Model-2: increased disparity>



<Model-3: decreased disparity>

In contrast, the third model would expect the opposite; the reduction in sentence length would be mainly observed among male defendants, as female defendants had received lenient treatment even before the Booker and Gall decisions.

Given the argument that the increased judicial discretion following Booker and Gall would lead to the increased extra-legal disparity in sentencing outcome (Bowmen, 2012; Richter, 2008) and that the judges should utilize this opportunity to ameliorate the rigid stance on gender equality in the federal district courts (Raeder, 2006), I expect that the reduction in overall sentencing severity observed in the wake of the Booker and Gall decision will be more pronounced with the female defendants in the form of two-way interactions, which leads to hypotheses 1 and 2.

Hypothesis 1: The impact of gender will differ between the pre-Booker and the post-Booker periods such that the sentence discount that female defendants receive will be greater in the post-Booker period than in the pre-Booker period.

Hypothesis 2: The impact of gender will differ between the pre-Booker and the post-Gall periods such that the sentence discount that female defendants receive will be greater in the post-Gall period than in the pre-Booker period.

The second group of hypotheses attempts to revisit the core question that this dissertation aims to address—that is, the effect of the ecological context of the court community. According to the theoretical perspectives discussed earlier, it is expected that the potential interaction effects between gender and Booker/Gall, if there is any, will also vary across different ecological contexts in the form of three-way cross-level interactions.

More specifically, I expect that female defendants would enjoy less benefit of reduced sentence severity following the Booker and Gall decisions in a district court characterized by more conservative political/religious climates. This expectation is based on a series of assumptions that, even though it is expected that there would be an overall reduction in sentence severity for women defendants following Booker/Gall (Hypotheses 1 and 2), the degree to which the female defendants receive the gender discount would depend on the characteristics of the federal court communities; the court communities characterized by higher levels of political/religious conservatism would give less of a discount to female defendants, as the informal sentencing norms entrenched in these court communities would not be as sympathetic to female defendants as other court communities characterized by lower levels of political or religious conservatisms. In addition, it may also be true that the impacts of Booker and Gall would be smaller in the districts with higher levels of conservatism, because the proportion that mandatory minimums account for would be greater in these district courts.

Hypothesis 3: Female defendants will enjoy less benefit of reduced sentence severity following Booker or Gall in a district court characterized by higher levels of religious conservatism.

Hypothesis 4: Female defendants will enjoy less benefit of reduced sentence severity following Booker or Gall in a district court characterized by higher levels of political conservatism.

CHAPTER 4: METHODOLOGY

DATA

This dissertation employs the USSC's Monitoring of Federal Criminal Sentences data files as a main data source on offenders convicted in U.S. District Courts. For the first study, I will use the data from FY2008 to FY2010 and, for the second study, will utilize data ranging from FY2001 to FY2010. The USSC data have been used by a number of studies (See, e.g., Johnson & Betsinger, 2009; Johnson et al., 2008) and have several strengths for this dissertation. For example, they contain detailed information that may partially explain lenient sentences afforded to female defendants, including whether a defendant plays a minor/minimal role in the crime and whether a defendant shows remorse by accepting responsibility.

Another important strength of the USSC data is that they include felony cases sentenced in all of the US District Courts, which provides a varying degree of social contexts within a single legal structure. The main USSC data set is also complemented by some contextual level data, which were obtained from several sources. To begin, court religious environment data were drawn from the Religious Congregation and Membership Study (American Statistical Association of Religious Bodies, 2002), and the County Characteristics data (ICPSR, 20660) was the source for court political environment. The 2000 U.S. Census data were also used to capture district-level variations in social structural characteristics (e.g., % African Americans). In addition, district-level index crime rates were obtained from the Uniform Crime Reports (UCR).

Lastly, information on federal judges was extracted from the Federal Judicial Center biographical data.

Dependent Variable

Following a one-stage model in judicial decision-making, I will employ one dependent variable throughout the dissertation: length of the prison sentence. Consistent with the prior studies examining sentence severity in federal courts (Kim et al., forthcoming; Schanzenbach & Tiller, 2007; Starr & Rehavi, 2013; USSC, 2010), all the probation and other alternative cases are included as zero months of incarceration.⁷ Length of prison sentence, originally measured in months, will be capped at 470 months, and log-transformed with the constant of one added due to extreme skew (See Johnson, 2006; Johnson & Betsinger, 2009).

Key Independent Variables

In accordance with the hypotheses put forward previously, the first key independent variable will be gender, which is measured as a dummy variable (female=1, male=0). At the district level, there are two contextual-level independent variables. Compiled from the 2000 Religious Congregation and Membership Study, religious conservatism is operationalized by the number of adherents to Conservative Protestant denominations per 1,000 persons in the population of the counties comprising the district where the court is located (see Ulmer et al., 2008; Vieratis et al., 2007). It should be

⁷ The original variable representing this variable is SENTTOT0, not SENTTOT.

noted that there are more than one way of classifying Protestant denominations. For the present study, the coded variable name in the original data set is evangelical Protestant, which includes fundamentalist and evangelical Protestants, but excludes Mormons.

Political conservatism was measured in a couple of different ways in previous studies. One of the most popular measures employed is the percentage of self-identified Republicans; that is, registered Republican voters (Jacobs & Helms, 1996). Even though this measure was heralded as the most effective indicator of political contexts by some scholars (see Shively, 1980), criticisms were also levied against its use (see Helms, 2009). For that reason, the present study employs actual voting patterns in relation to Republican Party affiliation, thus measuring political conservatism as the percentage of votes for George W. Bush in the 2004 presidential election, which was compiled from county characteristic data (ICPSR, 20660). Both the variables, originally measured at county level, are aggregated at the federal district level and were standardized (mean=0, SD=1), with higher scores indicating a higher level of religious or political conservatism. The scales are standardized because of the need to center measures that are used to create multiplicative interaction terms (Aiken & West, 1991), as well as to facilitate the easy interpretation concerning the interactions.

Control Variables

To address the concern that gender disparity may be at least in part a statistical artifact resulting from the lack of adequate control variables (Steffensmeier et al., 1993), a variety of control variables, at the defendant and district levels, will be included in the

analyses. Starting off with the individual-level controls, racial/ethnic minority status is captured with three dummy variables identifying Black, Hispanic, and other races, with whites serving as the reference category. The offender's number of dependent children is measured with four dummy variables, indicating if the defendant is responsible for one dependent child, two dependent children, three dependent children, and four or more dependent children, with no children serving as the reference category.⁸ Both linear and quadratic terms of the age of the offender are also included to take into account the possible non-linear relationship between age and sentencing severity (Steffensmeier et al., 1993). I will also include a control for education, captured by a dummy variable (some college or above=1), and citizenship status, separating U.S. citizens from non-U.S. citizens (non-U.S. citizen=1).

Prior research consistently shows that offenders' offense severity and criminal history scores are associated with sentencing outcomes. In line with prior works, the presumptive guidelines sentence will be employed, which combines the 43-point offense severity scale with the 6-point criminal history scale and accounts for sentencing adjustments that affect the final sentence under the federal guidelines.⁹ Consistent with

⁸ Given the importance of the dependent child status variable, a set of dummy variables is used to capture a potential non-linear effect. The decision to employ dummy variables to account for the non-linearity is justified under the assumption that there is no arbitrary cut-off issue with regard to this variable.

⁹ As noted earlier, variables indicating whether a defendant plays a minor/minimal role in the current crime and whether a defendant shows remorse, which is usually captured by the acceptance of responsibility discount, are available in the USSC data set. Since they are already built into the final presumptive sentence, I will not control for those two variables in the analysis.

the measurement strategy on the dependent variable, the presumptive sentence will be capped at 470 months and log transformed with the constant of one added. As will be explained in a greater detail, I will employ two different presumptive sentence measures for STUDY-2: one derived from the GLMIN variable and the other one from the XMINSOR variable from the federal monitoring data set. The main difference between the two presumptive sentences lies in the fact that the former takes into account the mandatory minimum trump while the latter does not. Thus by employing the latter presumptive sentence, I will be in a better position to estimate the impacts of whatever the Booker and Gall decisions entail, taking into account the fact that the application of mandatory minimums has also changed in the wake of Booker and Gall.

I will include additional controls for whether a defendant has criminal history (1=yes; 0=no), not the offender's final criminal history score,¹⁰ and a dummy variable capturing whether a case involves multiple counts (1=yes; 0=no). Four dummy variables are also controlled to capture the most serious offense type for which the offender was convicted: drug, fraud, firearm, and other offenses, holding violent offense as the reference category. Prior research also emphasizes that plea bargaining and pre-trial

¹⁰ This methodological choice also needs a brief mentioning. Some previous studies in federal sentencing employed both the offenders' final criminal history scores and the presumptive sentence (see Albonetti, 2002; Johnson et al., 2008). As noted, the final criminal history score is one of the two main determinants of the presumptive sentence. Therefore, this practice may do more harm than good. In this dissertation, I will employ a variable measuring whether a defendant has a criminal history. Given the way in which the final criminal history score is structured, it is important to control for this variable, as the criminal history category 1 does not differentiate a case with a zero point for a case with one point.

detention may affect sentencing severity (e.g., Albonetti, 1991; Ulmer & Johnson, 2004). Thus, I will control for these factors: plea bargaining, which distinguishes between guilty pleas and trial convictions (1=conviction resulting from plea bargaining; 0=otherwise), and pre-sentence release (1=in custody; 0=otherwise). Finally, I will include three dummy variables for departures from the guidelines: upward, substantial assistance, and downward departure, holding no departure as the reference category.¹¹

With regard to the court environment control variables, at the court organizational level, I will control for a set of court-level covariates. First, because prior research documented that judicial characteristics, even used in aggregated forms, affect sentencing outcomes in federal courts (Schanzenbach, 2005; Schanzenbach & Tiller, 2007), I will employ three measures for percent Republican presidential appointees, percent male judges¹², and percent white judges. Second, I will also incorporate caseload, measured as the average number of cases processed in a district court for a given year divided by the number of active judges. Third, because court size is considered one of the most important organizational variables (Eisenstein et al., 1988), court size is also controlled, which is measured by the number of authorized judgeships in each federal district.

¹¹ Throughout this dissertation, I will present two results with and without the departure status variables, as there is a debate on whether one should control for this important variable.

¹² Given a potential importance of gender composition of the judiciary on gender disparity (Kruttschnitt & Savolainen, 2009), the control of the aggregated measure of female judges will be also employed.

Table-1 Variables Description: Codes and Summary

Variable	Coding Scheme	Description
Sentence Length (DV)	Log (months)	Natural log of the total number of months of incarceration
Gender	1=female 0=male	Dummy indicator for sex
Race/ethnicity	4 dummy variables (Black, Hispanic, Others and white)	Dummy indicators for offender race/ethnicity with white as the reference
Age	Years	Continuous measure of age of defendant at the time of sentencing
Education	1=some college or above 0=high school or below	Dummy indicator for educational attainment
Children	4 dummy variables (None, one, two, three and four above)	Dummy indicators for dependent children with no children as the reference category
Citizenship	1=non-U.S. citizen 0=U.S. citizen	Dummy indicator for citizenship status
Detention	1=detained 0=not detained	Dummy indicator for offender's presentence detention status
Plea	1=pled guilty 0=pled not guilty	Dummy indicator for defendants who pled guilty
Multiple Counts	1=multiple counts 0=single counts	Dummy indicator for the presence of more than two counts
Criminal History	1=Yes 0=No	Dummy indicator for the presence of criminal history
Crime types	6 dummy variables (Drug, Fraud, Firearms, Property, others and violence)	Dummy indicators for the most serious offense type with the violence as the reference category
Departure Status	4 dummy variables (Within range, upward, SA, and other downward departures)	Dummy indicators for the departure status with the within range as the reference category
Presumptive Sentence	Log(months)	Natural log of minimum months of incarceration indicated by the guidelines
Conservative religion	Standardized	Continuous measure of the number of adherents to Conservative Protestant denomination among 1,000 county population, aggregated at the district level
Political Conservatism	Standardized	Continuous measure of the percentage of votes for George W. Bush in the 2004 presidential election at the district level
%Republican judges	Standardized	Continuous measure of the percentage of judges appointed by Republican president
%Male judges	Standardized	Continuous measure of the percentage of male judges within a district court
%white judges	Standardized	Continuous measure of the percentage of white judges within a district court
Caseload	Standardized	Continuous measure of cases process in a district divided by the number of judges
Court size	Standardized	Continuous measure of the number of authorized judgeships in a district court
% Black residents	Standardized	Continuous measure of the percentage of residents who are African Americans
Disadvantage	Standardized factor score	Continuous measure of socio-economic disadvantage
Crime rates	Standardized	Continuous measure of crime rates per 1,000 county residents, aggregated at the district level

Turning to the district environmental level, the following three covariates will be employed: First, I will control for racial composition of federal district, which measures a district's percent Black population. Second, a measure for socio-economic disadvantage is computed using a standardized factor score derived from four items extracted from the 2000 Census: percentage female-headed families with children, male unemployment rate, poverty rate, and the percentage of people without a high school diploma

(Eigenvalue=2.867, Factor loadings: minimum .722 and maximum .929, Alpha=.85). Third, crime rate will be also controlled (per 1,000 residents in district). Similar to Johnson and colleagues (2008), crime rate will be calculated by dividing the number of the UCR index crime by a district's total population, which is further multiplied by 1,000. I will grand-mean center all the control variables at the individual level, and standardize all the covariates at the district level for the convenience of interpretation. Brief information on the coding schemes and variable descriptions is provided in Table 1.

Common Analytic Issues

Throughout the dissertation, the following analytic issues will be addressed as follows:

Two-stage v. One-stage model

A recurring theme in empirical sentencing research is the correct way to model judicial decision-making on sentencing outcomes (see a recent debate for example: Ulmer et al., 2011b v. USSC 2010). In sentencing research, a more conventional

approach is to estimate two separate equations to model judicial decision-making: one for incarceration and the other for sentence length (Steffensmeier et al., 1993; Wheeler, Weisburd, & Bode, 1982).¹³ This approach, oftentimes referred to as a two-stage model, is premised on the assumption that judges make fundamentally different decisions when deciding whether to incarcerate and for how long (Fischman & Schanzenbach, 2012). In contrast, some scholars argue that judicial decision-making, especially under sentencing guidelines schemes, is better modeled utilizing a so-called one-stage model, where the outcome is only sentence lengths and the non-incarceration cases are included as a prison sentence of zero (Bushway & Piehl, 2000).

The scholars in favor of this latter approach argue that similar criteria guide judicial decision-making that influence both the in/out and sentence length decisions, as judges are instructed to first calculate a sentencing range, followed by a specific prison term under guidelines schemes (Bushway & Piehl, 2000; Starr & Rehavi, 2013; see also Paternoster, 2011). Meanwhile, this methodological choice also has some statistical implications. The main objection to the one-stage approach is the potential possibility that including probation cases as a zero month sentence and estimating an OLS model may yield a non-normal distribution of error, due to the extreme positive skew resulting from the zero month prison cases (see Ulmer et al., 2011b). Conversely, the camp advocating for the use of one-stage model criticizes the two-step approach for the lack of

¹³ It is worth noting that Holleran and Spohn (2004) further demonstrated that the so-called total incarceration variable, which combines prison and jail into a single outcome, should be reconsidered, as the jail and prison sentences are two separate decision-making processes. Instead they argued for the importance of employing a trichotomous outcome—probation, jail, and prison.

attention to a possible sample selection bias issue, as the two-stage approach is vulnerable to selection bias (Starr & Rehavi, 2013). Even though there exists some important breakthroughs to handle the selection issue (see Berk, 1983; Bushway, Johnson, & Slocum, 2007; Heckman, 1976), it is extremely difficult, though not impossible, to address this issue in a methodologically correct manner and, in many cases, it is more likely the case that the sample selection issue is simply sidestepped (Bushway & Piehl, 2000; see Bushway et al., 2007 for an exception).

Yet another related issue concerns the appropriate way to address the zero month cases included under the one-stage model. Many legal scholars trained in econometrics claim that an OLS regression would approximately estimate an average coefficient of interest with robust standard errors being employed, as the deflated standard error issue resulting from a non-normal error term would be addressed by employing robust standard errors (see Fischman & Schanzenbach, 2012; Rehavi & Starr, 2012). From a slightly different angle, some prefer to employ a Tobit model based on the assumption that the probation cases are censored, because the variation in the severity of probation is not directly observed (see Albonetti, 1997; Bushway & Piehl, 2000). Upon a closer inspection, however, the sentence length is not censored at zero, rather non-incarceration sentences are known zeros, not any unknown latent variable (Starr & Rehavi, 2013). In addition, the Tobit model is known to be vulnerable to the violations of normality and homoskedascity assumptions (Cameron & Trivedi, 2010), which is more likely the case in this specific situation. Thus in this dissertation, I will employ the one-stage model as a

main analytic framework along with robust standard errors to address the potentially deflated standard errors.

Departure and mandatory minimum status as controls

Unlike the states' counterparts, in federal sentencing research, it is by now an almost established practice that one controls for the offender's departure status in both in/out and sentence lengths models (see for example, Albonetti, 1997; Doerner & Demuth, 2010; Feldmeyer & Ulmer, 2009; Kautt, 2000; Ulmer et al., 2010). A review of the prior studies, however, suggests that there seems to be no clear reason behind this methodological choice and it is hard to find even a passing comment on why this has been done this way. But a speculation is that the departure status variable, considered a legally relevant variable by many researchers and especially the USSC, would yield a robust estimate of interest, when used as a control variable. That is, the more control variables in a model, the more precise the estimates would be (See, for example USSC, 2004, P.D-11). This issue did not receive serious thought until a group of legal scholars criticized this practice (See Fischman & Schanzenbach, 2011, 2012; Freeborn & Hartmann, 2010; Rehavi & Starr, 2012). For instance, Starr and Rehavi (2013, p. 20) attempted to lay out the detailed ramifications resulting from the methodological choice by referring to a specific example of racial disparity published in USSC's Booker report. According to the authors,

“In effect, the Commission is estimating race gaps in the size of departures (and in the sentence choices within the narrow Guidelines range), but filtering out whether there is a departure and, if so, in what direction. That is, to say the least, *a strange choice*, and it could very easily produce misleading results.....And there is no apparent substantive

reason that differences in departure rates should be ignored when assessing sentencing disparities.” (Starr & Rehavi, 2013, p. 20)

This issue takes on added importance with regard to estimating the impact of Booker/Gall in federal sentencing, as the departures are unarguably the main channels by which the effects of Booker/Gall influence judicial decision-making. Therefore, controlling for a departure status is similar to curtailing the sources of changes when assessing the impacts that the Booker/Gall decisions have (see Engen, 2009; Paternoster, 2011).

The practice for departure control especially under the Booker/Gall context is also criticized for an endogeneity problem, as the departure status is impacted by the Booker/Gall decision as much as the sentence length and incarceration decisions are. In a similar context, the use of mandatory minimums as a control variable also creates the same problem. Some prior studies discovered that federal prosecutors have attempted to ameliorate the impact of the increased judicial discretion in the wake of Booker and Gall by invoking more mandatory minimums (see Fischman & Schanzenbach, 2012; Starr & Rehavi, 2013). Then mandatory minimums are clearly endogenous to the sentencing outcome and controlling for the application of mandatory minimums would produce biased estimates especially under the Booker/Gall context. In order to address this issue, I will employ a presumptive sentence variable derived from the XMINSOR, not GLMIN, from the official federal sentencing monitoring data set. The XMINSOR variable is based on an original sentencing range determined under the guidelines, which is not constrained by the mandatory minimums. Given this potential methodological debate, in this

dissertation, I will provide a series of supplemental models to give insights on to what extent these methodological choices would produce results that may be different from the ones reported in the previous studies.

Missing Data

Missing data is another complex statistical issue, as the issue could potentially damage both the validity and generalizability of a study (Allison, 2000). Given the importance of this issue, there are a variety of approaches to dealing with missing data problems, some of which include a replacing missing values approach (King, Honaker, Joseph & Scheve, 2001), a listwise deletion approach (Little, 1992), and a multiple imputation approach (Allison, 2000). Consistent with the previous studies, this dissertation will employ the listwise deletion approach for the following reasons.

First, this approach is consistent with all the previous studies in federal sentencing (see, for example Albonetti, 1997; Doerner & Demuth, 2010; Feldmeyer & Ulmer, 2009; Kautt, 2000; USSC, 2012). To my knowledge, in federal sentencing, there is no study which employs a different approach. Second, more importantly, listwise deletion is argued to perform better than other alternatives, including for instance, multiple imputation under some circumstances. To begin, it is suggested that listwise deletion would still yield valid inferences when the missing data problem is prevalent with the regressors, not the outcome variables (Little, 1992). And it has been generally known that The Federal Sentencing Monitoring Data set is well compiled with the missing data issue being relatively minor. For instance, in the present data files, fewer than 2% of cases have

missing values on such key variables as gender and sentencing outcomes and most variables in the data set have missing values fewer than 5%. In addition, the listwise deletion approach is less vulnerable from the violation of certain assumption underlying the most promising alternative-multiple imputation, such as MAR (Missing At Random).

STUDY-1

Data

The original USSC FY2008—FY2010 data file included 241,796 offenders sentenced in 94 district courts. Consistent with prior research (Johnson et al., 2008; Ulmer et al., 2011a), some cases will be removed from the analysis, including all immigration cases (N=75,860) and cases handled in Puerto Rico, Guam, Virgin Islands, the District of Columbia, and Northern-Mariana Island (N=3,395). After further removing cases sentenced prior to the U.S. v. Gall/Kimbrough decision from FY2008 (N=9,472) to control for any variation resulting from the decision, the final sample consists of 153,069 defendants sentenced in 89 district courts.

Measures

I will employ all the variables presented in the measurement section, including one dependent variable, two independent variables at the district level, along with all the control variables. The only variable added to the current analysis is a time variable. I will

include dummy variables for sentencing year to control for potential time trends, holding 2008 as the reference year.

Analytic Strategy

Given the nested nature of data—that is, individual cases are nested within 89 federal district courts—and the main interest in estimating cross-level interactions, multilevel modeling will be employed as the main analytic tool (Raudenbush & Bryk, 2002). The importance of utilizing multilevel modeling is well established in the sentencing literature. To illustrate, multilevel modeling allows researchers to estimate corrected standard errors, model the heterogeneity in regression coefficients and, most importantly to the present study, properly estimate cross-level interaction effects (see Britt, 2000; Spohn & Fornango, 2009; Wooldredge, 2010). I will analyze the data using STATA-13, and present model estimates with robust standard errors to account for the abnormal distribution of errors. In order to answer the research questions, the following two-level multilevel models will be estimated as a baseline model.

$$Y_{ij} = \beta_{0j} + \beta_1 \mathbf{Female} + \beta_2 \mathbf{X}_{ij} + e_{ij} \quad (1) \text{ Level 1}$$

$$\beta_{0j} = \gamma_0 + \gamma_{01} \mathbf{Politics} + \gamma_{02} \mathbf{Religion} + \gamma_{0j} \mathbf{W}_j + u_{0j} \quad (2) \text{ Level 2}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \mathbf{Politics} + \gamma_{12} \mathbf{Religion} + \gamma_{1j} \mathbf{W}_j + u_{1j} \quad (3) \text{ Level 2}$$

$$\beta_{2j} = \gamma_{20} \quad (4) \text{ Level 2}$$

Equation (1) is the primary model for the sentence length decision where Y_{ij} is the observed sentence length for defendant i in district j . β_1 represents the effect of gender on sentence lengths and X_{ij} denotes a vector of the characteristics of an individual defendant used as control variables. Equation (2) through (4) represent the level-2 analysis. Equation (2) model the intercept-the average sentence length for district j , as a function of level-2 variables. Thus, the coefficients for γ_{01} and γ_{02} measure the main effects of political and religious conservatism on the outcome respectively. W_j denotes a vector of district-level control variables in district j . Equation (3) plugged into equation (1) concerns the main research question on the cross-level interaction between gender and political/religious conservatism. Finally, u_{0j} and u_{1j} are random effects for the district and the gender, assumed to be normally distributed with means of zero and the variance of σ^2_μ and σ^2_ν .

In order to answer the 6 hypotheses explained previously, the analyses will proceed as follows: First, I will estimate an unconditional model to assess if there is significant variation in sentence length across districts. Second, a full multi-level model will be estimated to see whether gender has an effect on the outcome. Then, I will assess if gender-based sentencing disparity is moderated by religious and political conservatism by estimating a series of multilevel models that include cross-level interactions. Finally, to investigate to what extent the potential moderating effects of religious and political conservatism are further colored by racial and ethnic lines, I will partition the data by race/ethnicity and examine if the moderating effects of religious and political conservatism are stronger for white versus minority defendants. In doing so, I will also

perform a series of Z-tests (Paternoster et al., 1998) to demonstrate that the potential cross-level interactions between gender and religious/political conservatism significantly differ by race/ethnicity of defendants.

STUDY-2

Data

As noted previously, the second study utilizes the USSC standard research data set from FY2001 to FY2010, along with the 2000 U.S. Census, Uniform Crime Reports (UCR) data, biographical data on federal judges from the Federal Judicial Center, and the County Characteristics data (ICPSR 20660). The original USSC data file included 724,297 offenders in 94 district courts, but, as in line with the justification presented in the STUDY-1 data section, I will exclude some cases, including all the immigration cases and cases adjudicated in the U.S. territories. Finally, I will delete cases from the first quarter of FY2001 to account for a temporal order issue with the crime rate variable. With these cases being excluded, the final data file includes 519,015 offenders who were sentenced in 89 district courts.

Measures

Consistent with the explanation presented in the previous section, I will employ most of the variables in the common measurement section, but will also include some other important variables in this study. One of the most important variables only utilized

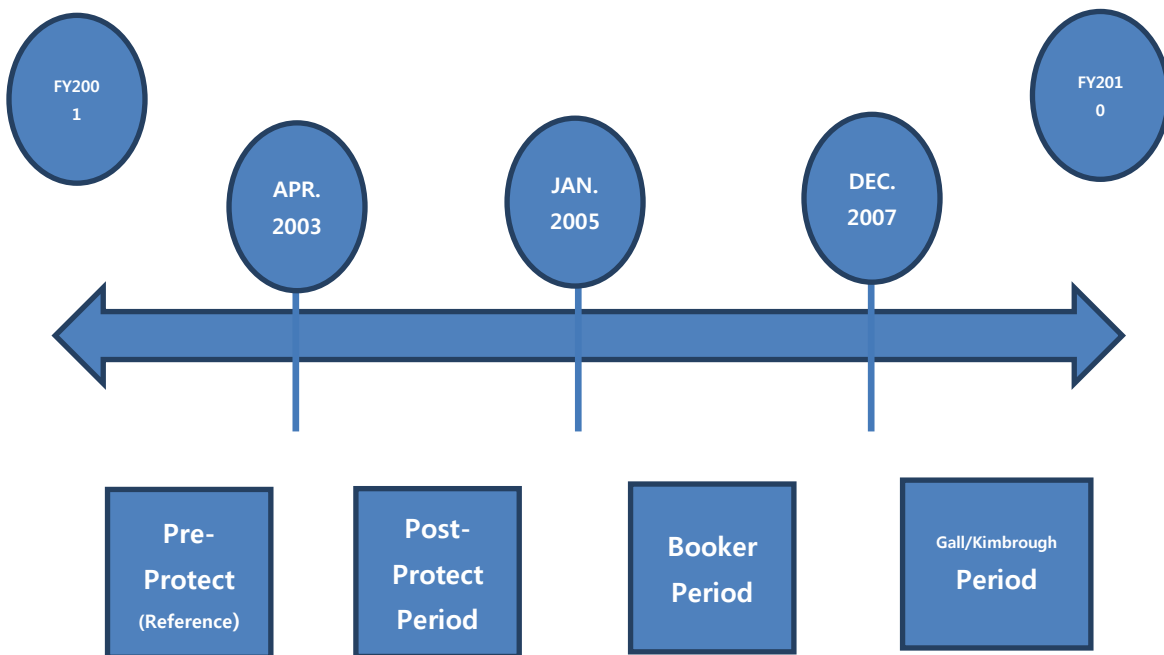
in the current study is a series of binary indicators representing the Booker and Gall decisions. As the focus of STUDY-2 is on comparing sentencing patterns across the distinct time periods and the data portioning approach taken by Ulmer et al. (2011a; 2011b), and USSC (2006; 2012) has some critical limitations, I will move away from the approach and elect to use all the cases in one model.

Therefore, constructing Booker/Gall variables is of utmost importance: the Booker variable will measure whether the cases were sentenced after the Booker decision was handed down, which will be coded as 1 if the offender was sentenced from the time the Booker ruling was announced to the time the Gall decision was handed down and as 0 otherwise. Similarly, the Gall variable will be coded as 1 if the case was sentenced after the Gall decision and as 0 otherwise. Regarding the use of reference category, I will use the pre-Protect era in order to take into account the unique effect that the PROTECT Act had (see Ulmer et al., 2011a, 2011b for the same approach). Thus, there are four binary time indicators: pre-Protect (reference category), post-Protect, Booker, and Gall as presented Figure-5 below.

Control variables at the individual level are the same as presented in the STUDY-1, except for the following two variables. Due to a complex modeling strategy employed for the current study, I will use the dependent children variable measured as a binary dummy variable (1=having dependents, 0=no dependents). In addition, the age variable will be also included without the squared term. At the district level, I will also employ the same variables, such as judicial characteristics, district caseloads, and district

crime rates, but in different forms. Unlike the situation for the STUDY-1, these variables will be used as time-varying covariates to rule out the possibility that any concurrent factors, other than Booker and Gall, would influence the changes in sentence severity over time. In a similar context, a time variable will be also incorporated—measured as the month of sentencing, including a time squared term—to account for any unmeasured and gradual time trends in sentencing outcomes (See Starr & Rehavi, 2013).

Figure-5 Booker and Gall Variables Coding



Lastly, I also control for the three time-invariant court-level variables, which were presented above: disadvantage, % Black residents, and court size. Finally, given the recent debate on whether departures status should be controlled in a model, especially in

Booker and Gall contexts (see Ulmer et al., 2011 v. Starr & Rehavi, 2013), I will provide two different results from two different models, one with departures and the other without departures.

Analytic Strategy

As suggested, I will employ models in which all the cases are included for the four distinct periods. Previous studies on Booker and Gall, especially Ulmer et al. (2011a; 2011b), built their entire analyses on the data partitioning approach, which may be described as an “easy to perform and understand” strategy. However, this approach has its own limitations. First, the partitioning technique may yield an erroneous result, as the approach basically estimated four different models with their own unique error structures, instead of having one common error structure. Under some circumstances, making a comparison across different models with different error structures may lead to imprecise estimates. Second, the partitioning approach does not go well with a temporal context issue (see Ulmer & Light, 2013). Especially when the data partitioning is premised on the long-term time periods, it is becoming difficult to address the long-term time trend issue and to control for time varying covariates, as this may have been the case with the Ulmer et al. (2011a; 2011b) and USSC (2006; 2010) studies. Finally, the partitioning approach is also not in a good position to quantify the overall impact of Booker and Gall, as the focus of this approach is more directed toward making a comparison between the same variables across different time periods. As suggested, however, calculating the overall

Booker/Gall effects would be critical, especially if one wants to pinpoint the actual amount of changes in gender disparity that occurred in the wake of Booker and Gall.

For this specific study, I will use a repeated cross-sectional design (see Menard, 2002) and employ an analytic strategy using a three-level hierarchical linear modeling, in which individual cases are nested within time and districts (see DiPrete & Grusky, 1990; Kim et al., forthcoming; Xie, Lauristen, & Heimer, 2012). Because it is plausible that a defendant sentenced in a similar time point shares similar sentencing trends, that is, individual cases are nested within time, and a defendant sentenced in the same district is treated in a similar manner, that is individual cases are nested with places, this unique data nesting issues should be taken into account when estimating models. In order to provide the modeling framework, I will use the following model as a baseline model.

$$Y_{ij} = \beta_{0j} + \beta_1 Female + \beta_2 \mathbf{X}_{ij} + e_{ij} \quad (1) \text{ Level 1}$$

$$\beta_{0j} = \pi_{0j} + \pi_{01j} Booker_{tj} + \pi_{02j} Gall_{tj} + \pi_{03j} Time/Time^2_{tj} + \pi_{04j} \mathbf{Z}_{tj} + r_{tj} \quad (2) \text{ Level 2}$$

$$\beta_1 = \pi_{10j} + \pi_{11j} Booker_{tj} + \pi_{12j} Gall_{tj} + \pi_{13j} Time/Time^2_{tj} + \pi_{14j} \mathbf{Z}_{tj} + r_{tj} \quad (3) \text{ Level 2}$$

$$\pi_{0j} = \gamma_{00} + \gamma_{01j} Politics + \gamma_{02j} Religion + \gamma_{03j} \mathbf{W} + u_{0j} \quad (4) \text{ Level 3}$$

$$\pi_{10j} = \gamma_{10} + \gamma_{11j} Politics + \gamma_{12j} Religion + \gamma_{13j} \mathbf{W} + u_{10j} \quad (5) \text{ Level 3}$$

Starting off with the equation (1), where level-1 variables are specified, Y_{ij} is the observed sentence length for defendant i , at time t , in district j . β_1 measures the impact of gender on sentence lengths. \mathbf{X}_{ij} denotes a vector of the characteristics of an individual

defendant i , at time t , in district j . At level 2, where district level time varying variables are employed, it is important to note that there is no individual defendant i subscript in any of the level-2 variables. The Booker and Gall dummy variables are also important predictors of interest. \mathbf{Z}_{ij} is a vector of time-varying covariates at the district level, such as crime rate, caseload, and aggregated judicial characteristics measured at time t in district j . \mathbf{Time}_{ij} is measured as the month when cases were sentenced in order to capture any long-term time trend effects. At level 3, where a subscript t is not found, \mathbf{W}_j denotes a vector of time-invariant covariates in district j , such as the district level independent variables, such as political and religious conservatism, along with district level time-invariant control variables. Finally, \mathbf{r}_{ij} and \mathbf{u}_{0j} are random effects for the time and the district, which are assumed to be normally distributed with means of zero and variances of σ^2_{μ} and σ^2_{ν} .

Against this backdrop of the current modeling framework, for the first two hypotheses, the main interest lies in estimating two-way, cross-level interactions between Female and Booker/Gall to investigate whether the effect of gender is larger or smaller during the Booker/Gall period, as compared to the pre-Protect period. These questions will be answered by the results from the models estimated through equation (1) in conjunction with (3). In order to answer the latter two hypotheses, I will estimate three-way cross-level interactions among Female (level-1), Booker/Gall (level-2), and Politics/Religion (level-3) to examine whether the potential changes in gender effect across the Supreme Court decisions have distinctive impacts depending on the level of court communities' political and religious conservatism.

CHAPTER 5: ANALYSES AND RESULTS

Chapter 5 provides the findings for STUDY-1 and STUDY-2. Each study starts with descriptive statistics to present a description of study subjects and correlation matrices to examine bivariate relationships among variables. In order to provide specific tests for the hypotheses presented in the previous chapter, findings from a series of multivariate analyses will be presented. Finally, each study concludes with a summary of main findings.

STUDY-1

Main Hypotheses

Hypothesis 1: Female defendants will receive shorter sentences than their male counterparts.

Hypothesis 2: Female defendants adjudicated in jurisdictions with higher levels of religious conservatism will receive smaller sentence discounts relative to other female defendants adjudicated in jurisdictions with lower levels of religious conservatism.

Hypothesis 3: Female defendants adjudicated in jurisdictions with higher levels of political conservatism will receive smaller sentence discounts relative to other female defendants adjudicated in jurisdictions with lower levels of political conservatism.

Hypothesis 4: Gender disparity will be least pronounced in jurisdictions with higher levels of both religious and political conservatism.

Hypothesis 5: The interaction effect between female and religious conservatism—the degree to which religious conservatism reduces the female discount—will be significantly greater for non-white female offenders, as compared to white female offenders.

Hypothesis 6: The interaction effect between female and political conservatism—the degree to which political conservatism reduces the female discount—will be significantly greater for non-white female offenders, as compared to white female offenders.

Descriptive Statistics

The descriptive statistics for all the variables are presented in Table-2. Starting off with the dependent variable, the length of the prison sentence has a mean of 3.13 and a standard deviation of 1.72 with approximately 1.5% of cases with missing values.

Moving on to the key independent variables of the current study, the majority of defendants were males, with the female defendants accounting for only about 15% of the sample. The descriptive statistics for the two contextual-level independent variables also indicated a substantial level of variation among federal district courts; conservative religion, measured as the number of adherents to Conservative Protestant denominations per 1,000 persons, has a mean of 188 and a standard deviation of 132. The political conservatism variable, represented as the percentage of votes for George W. Bush in the 2004 presidential election, has a mean of 57% and a standard deviation of 8.6% with the minimum of 34.5% and the maximum of 77.45%.

Turning to the descriptive statistics for control variables, the racial/ethnic composition of the sample indicated that the offenders were relatively equally distributed

across the racial/ethnic categories with whites accounting for 36%, African Americans representing 30%, Hispanic occupying 27%, and others 7%. Roughly about 37% of defendants did not have dependent children while 63% of defendants had one or more dependents. On average, the offenders were predominantly U.S. citizens (79%), with a mean age of 35 and their average level of education being less than some college above (74%). Regarding legal variables of the defendants in the sample, a substantial number of defendants had a criminal history (76%), pled guilty (95%), and were held in custody prior to sentencing (68%). In addition, most offenders in the sample were convicted of drug offenses (45%), followed by fraud (19%), firearms (14%) and other offenses (12%). Finally, slightly more than half of defendants received within guidelines sentences (54%), with 18% of defendants receiving substantial assistance departures and 26% of defendants receiving other judicial departures.

Moving on to the contextual-level control variables, the descriptive statistics for the court organization variables showed that the federal cases were presided over by white (81%) and male judges (79%). It was also shown that more than half of the federal judges were appointed by Republican presidents (61%). Federal district courts also varied in terms of their sizes demonstrated by the mean value of 10.26 judges and the standard deviation of 6.51 judges per district. Caseload pressure, measured as the number of cases per year divided by the number of active judges, also indicated a substantial level of variability across federal districts, with a mean value of 170 cases per judge; the maximum was 630 cases and the minimum was 36 cases.

Table-2 Descriptive statistics

Dependent variable		Mean	SD	Min	Max
Sentence length (logged)		3.13	1.72	0	6.15
Independent and Control variables		Mean	SD	Min	Max
Level 1 Variables	Female	.15	.36	0	1
	White (reference)	.36	.48	0	1
	African American	.30	.46	0	1
	Hispanic	.27	.44	0	1
	Others	.04	.21	0	1
	No dependent (reference)	.37	.48	0	1
	One dependent	.19	.39	0	1
	Two dependents	.17	.38	0	1
	Three dependents	.12	.32	0	1
	Four and above dependents	.12	.33	0	1
	Age	35.70	11.51	16	90
	Some College above	.26	.43	0	1
	Non-citizen	.21	.41	0	1
	Plea	.95	.21	0	1
	Detained	.68	.46	0	1
	Violent (reference)	.04	.20	0	1
	Drug	.45	.49	0	1
	Fraud	.19	.39	0	1
	Firearms	.14	.35	0	1
	Property	.03	.17	0	1
	Others	.12	.36	0	1
	Multiple counts	.22	.41	0	1
	Criminal History	.76	.42	0	1
	Presumptive sentence (logged)	3.55	1.49	0	6.15
	Within guidelines (reference)	.54	.49	0	1
	Upward departure	.02	.13	0	1
	SA departure	.18	.38	0	1
	Downward departure	.26	.43	0	1
FY 2008 (reference)	.29	.45	0	1	
FY 2009	.35	.47	0	1	
FY 2010	.36	.47	0	1	
Level 2 variables	Conservative religion	188.17	132.71	14.01	517.28
	Political Conservatism	57.16	8.60	34.55	77.45
	% Black residents	9.32	10.36	.15	44.83
	Disadvantage (z)	0	1	-1.97	2.33
	Crime rates	10.66	12.39	.76	47.29
	% Republican judges	61.04	17.05	0	100
	% Male judges	79.40	12.05	50	100
	% White judges	81.12	14.70	47.05	100
	Caseload	179.2	160.9	36.3	630
	Court size	10.26	6.51	1	28

All the variables are presented as their original forms unless otherwise specified.

Federal districts also differed significantly in terms of the characteristics associated with court environmental factors. To start, the descriptive statistics for racial composition of district indicated that the average percent Black was 9.32%, with a maximum of 44.83% and a minimum of 0.15%. The socio-economic disadvantage of districts also varied from a maximum of 2.33 to a minimum of -1.97. The average crime rate is shown to be 10.66 per every 1,000 residents in a district, with a minimum of 0.76 and a maximum of 47.29.

Bivariate Relationships

Examining bivariate relationships allows for checking for any potential problematic collinearity issues before moving onto multivariate analyses. Here in STUDY-1, I provide two different correlation tables for Level-1 and Level-2 variables for ease of presentation. I will start with a table incorporating the zero-order correlations among individual-level measures. The first column of Table-3 indicated that most legal variables, such as presentencing detention status, multiple counts, and criminal history, were moderately correlated with sentence length, with the correlations ranging from 0.22 to 0.49. The only prominent exception was the correlation between presumptive sentence and sentence length ($r=-0.84$), which is consistent with and well documented in the sentencing literature. Regarding correlations for extra-legal variables, two variables stood out: gender ($r=-0.28$) and African Americans ($r=0.14$). Considered together, most of correlation among independent variables in Table-3 fell between 0.53 and -0.45, which was well short of creating a potential collinearity issue.

Table-3 Correlation Matrix for Individual-Level Variables

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	
V1	1																						
V2		1																					
V3			1																				
V4				1																			
V5					1																		
V6						1																	
V7							1																
V8								1															
V9									1														
V10										1													
V11											1												
V12												1											
V13													1										
V14														1									
V15															1								
V16																1							
V17																	1						
V18																		1					
V19																			1				
V20																				1			
V21																					1		
V22																						1	

V1(Sentence Length), V2(Female), V3(Black), V4(Hispanic), V5(Other), V6(1-kid), V7(2-kids), V8(3-kids), V9(4 or above-kids), V10 (age), V11(some college above), V12(non-Citizen), V13(detention), V14(Plea), V15(Presumptive sentence), V16(Multiple Count), V17(Criminal History), V18(Drug), V19(Fraud), V20(Firearms), V21(Other), V22(Property)

Turning to Table-4, which displays zero-order correlations among level-2 variables, the first row of Table-4 showed that religious conservatism was highly correlated with political conservatism ($r=0.55$) and disadvantage ($r=0.49$). It further indicated that % Black ($r=0.31$) and the size of court ($r=-0.36$) were also moderately correlated with the religious conservatism measure. Meanwhile, the political conservatism measure turned out not to be as highly correlated as the religious conservatism measure was correlated with other measures. More specifically, the second column of Table-3 suggested that the political conservatism measure had only moderate levels of correlations with the size of court ($r=-0.37$) and % male judges ($r=0.31$). Taken together, Table-4 suggested that most of correlations fell between -0.39 and 0.55 , which was higher correlations as compared to those of individual-level measures.

Table-4 Correlation Matrix for District-Level Variables

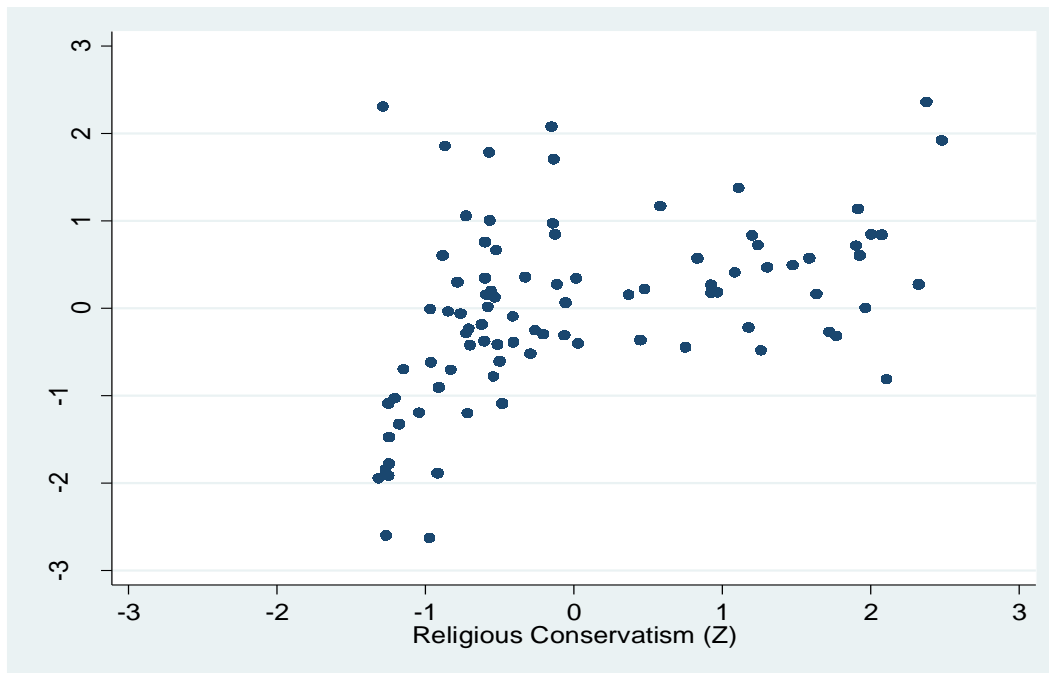
	V1	V2	V3	V4	V5	V6	V7	V8	V9
V1	1								
V2	.55	1							
V3	.49	.20	1						
V4	.31	-.15	.36	1					
V5	-.36	-.37	-.05	-.02	1				
V6	.06	.29	.46	-.18	.12	1			
V7	.04	-.05	.08	.13	.04	-.02	1		
V8	.21	.15	.19	.09	-.15	.17	0	1	
V9	.26	.31	.02	.04	-.39	.07	-.04	.26	1
V10	.08	.04	-.22	.11	-.51	-.61	-.04	.04	.17

V1(Religious Conservatism), V2(Political Conservatism), V3 (Disadvantage), V4 (%Black), V5 (Size), V6 (Caseload), V7 (Crime rates), V8 (% Republican Judges), V9 (% Male Judges), V10 (%White Judges)

As the relationship between religious and political conservatism is both theoretically and empirically important, I additionally present Figure-6, which displays a scatterplot for the two variables. One specific concern that could potentially arise is to what extent these two variables are distinct or similar, as there is a moderate correlation between these variables. Starting off with the distributions of each variable, it appeared that religious conservatism had limited variability especially around the lower end of its measure, while the political conservatism measure was evenly distributed. This scatterplot basically revealed a positive association between two measures, even though the association was getting weaker especially around the higher end of two variables. Finally, there were some cases which fell outside the overall pattern of the scatterplot. Especially noteworthy were two cases representing District of Utah and District of Arkansas East with the former having an extremely high value of political conservatism but a lower value of religious conservatism and the latter having the opposite pattern.¹⁴ Overall, the univariate and bivariate analyses presented thus far suggest that there were no signs of problematic collinearity that could potentially harm the estimation of multivariate models.

¹⁴ The fact that the district of Utah is located around the lower end of the religious conservatism measure may be surprising. This is because the conservative Protestant measure employed in the current study does not include Mormons, which accounts for a large proportion of population residing in Utah. As noted above, there exists a debate as to what kinds of religious groups belong to conservative Protestant (see Woodberry & Smith, 1998). However, some researchers found that Mormons had relatively less conservative gender attitudes, as compared to fundamentalists (Brinkerhoff, Jacob & Mackie, 1987).

Figure-6 Scatterplot for Religious and Political Conservatism



Unconditional Model

Before presenting substantive results regarding the specific hypotheses in study-1, one unconditional model was estimated to provide evidence that the ICCs are large enough to warrant multi-level modeling. The results, presented in Table-5, indicated that a statistically significant variation indeed existed in the average sentence length. To be more specific, the intercept (3.198) represented the grand-mean logged sentence length that offenders received in the sample. The estimate of the random variance across districts showed that there was a statically significant variation in the sentence lengths across districts ($SD=0.404$, $p<.001$), which meant that sentence lengths significantly differed across district courts. In order to give a more realistic idea of how the logged

sentence length varied across district, both the 95% of confidence interval¹⁵ and the Intra-Class Correlation (ICC)¹⁶ were calculated (see Rabe-Hesketh & Skrondal, 2008). The calculated 95% confidence interval turned out to fall between 2.40 and 3.98 of logged sentence length, which means that the average logged sentence length for 95% of the district courts ranged from 2.406 to 3.989. The calculated ICC was 0.055, which means that about 5.5% of sentence variations could be attributable to the district in which the sentence was imposed. Overall, the evidence provided in this section justifies the use of multi-level modeling to answer the specific research questions.

Table-5 Unconditional Model

Fixed Effects	Coefficients	SE	P
Intercept	3.198***	.043	.000
Random Effects	SD	SE	P
Level-2	.404***	.030	.000
Residual	1.678	.003	

*p <0.05, **p<0.01, ***p<0.001

Sentence Length Decision: Hypothesis-1 through -4

In order to address the relationships predicted by hypotheses 1 through 4, a series of random intercept models were estimated with the control variables added. The results are provided in Table-6, concerning the direct effect of gender (H-1), the interaction

¹⁵ The formula to calculate the range is $(3.198 \pm (1.96 * 0.404))$.

¹⁶ ICC could be conceptualized as correlations among cases within a group and is calculated, in this specific case, as the between district variance divided by the overall variance.

effects involving gender and religion (H-2), gender and politics (H-3), and the three-way interaction among gender, religion, and politics (H-4) related to sentence length decision.

Starting from the overall results, Model 1 in Table-6 suggested that the logged sentence length still varied across district courts, evidenced by the statistically significant random variance component associated with the intercept ($SD=0.133$), even after controlling for relevant covariates. This means that sentencing practices differed significantly across federal district courts independent of case compositional differences and some district-level covariates included in the model. The general patterns of the results were also largely congruent with those of previous studies. Most legal variables were statistically significantly associated with the sentence length decision in expected directions; first, the presumptive sentence and the presence of criminal history both increased logged sentence lengths, along with the detention status and the multiple counts variables. In regard to the results concerning extra-legal factors, African American, Hispanic, and other offenders received longer sentences, as compared to their white counterparts. Both citizenship status and educational attainment were also significantly associated with the logged sentence length, albeit in different directions. The effect of dependent child status on sentence length was consistently negative. Finally, the examination of the results on contextual-level covariates showed that both the crime rate and the court size variables were negatively associated with the logged sentence length.

Moving onto the specific results for each hypothesis, first, in line with the expectation on H-1, the results reported in Model-1 in Table-6 suggested a strong gender

Table-6 Sentence Length Models

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Female	-.257*** (.019)	-.260*** (.017)	-.258*** (.017)	-.256*** (.017)
Religion (z)	.023 (.019)	.021 (.019)	.023 (.019)	.029 (.019)
Politics (z)	.063*** (.017)	.063*** (.017)	.061*** (.017)	.057*** (.017)
Female*Religion	-	.062***(.016)	-	.043* (.016)
Female*Politics	-	-	.072***(.017)	.049* (.017)
Religion*Politics	-	-	-	-.009 (.017)
Female*Religion*Politics	-	-	-	-.008 (.017)
Intercept	3.295*** (.028)	3.295*** (.028)	3.295*** (.028)	3.300*** (.028)
Black	.106*** (.006)	.106*** (.006)	.106*** (.006)	.106*** (.006)
Hispanic	.089*** (.007)	.089*** (.007)	.089*** (.007)	.089*** (.007)
Others	.024* (.011)	.024* (.011)	.024* (.011)	.024* (.011)
One dependent	-.025*** (.006)	-.025*** (.006)	-.025*** (.006)	-.025*** (.006)
Two dependents	-.050*** (.006)	-.050*** (.006)	-.050*** (.006)	-.050*** (.006)
Three dependents	-.052*** (.007)	-.052*** (.007)	-.052*** (.007)	-.052*** (.007)
Four & above dependents	-.050*** (.007)	-.050*** (.007)	-.050*** (.007)	-.050*** (.007)
Age	.001*** (.000)	.001*** (.000)	.001*** (.000)	.001*** (.000)
Age ²	-.000*** (.000)	-.000*** (.000)	-.000*** (.000)	-.000*** (.000)
Education	-.064*** (.005)	-.064*** (.005)	-.064*** (.005)	-.064*** (.005)
Non-citizen	.092*** (.007)	.092*** (.007)	.092*** (.007)	.092*** (.007)
Plea	-.277*** (.010)	-.277*** (.010)	-.277*** (.010)	-.277*** (.010)
Detained	.641*** (.005)	.641*** (.005)	.641*** (.005)	.641*** (.005)
Drug	-.156*** (.011)	-.156*** (.011)	-.156*** (.011)	-.156*** (.011)
Fraud	-.193*** (.012)	-.193*** (.012)	-.193*** (.012)	-.193*** (.012)
Firearms	-.133*** (.012)	-.133*** (.012)	-.133*** (.012)	-.133*** (.012)
Property	-.202*** (.018)	-.202*** (.018)	-.202*** (.018)	-.202*** (.018)
Others	-.111*** (.012)	-.111*** (.012)	-.111*** (.012)	-.111*** (.012)
Multiple counts	.123*** (.005)	.123*** (.005)	.123*** (.005)	.123*** (.005)
Criminal history	.153*** (.006)	.153*** (.006)	.153*** (.006)	.153*** (.006)
Presumptive sentence	.863*** (.002)	.863*** (.002)	.863*** (.002)	.863*** (.002)
% Black residents (z)	.026 (.017)	.026 (.017)	.026 (.017)	.026 (.017)
Disadvantage (z)	.020 (.020)	.020 (.021)	.020 (.021)	.020 (.021)
Crime rates (z)	-.015 ⁺ (.009)	-.016 ⁺ (.009)	-.016 ⁺ (.009)	-.016 ⁺ (.009)
% Republican judges (z)	.002 (.012)	.002 (.012)	.002 (.012)	.002 (.012)
% Male judges (z)	.001 (.014)	.001 (.014)	.001 (.014)	.001 (.014)
% White judges (z)	-.000 (.020)	-.000 (.020)	-.000 (.020)	-.000 (.020)
Caseload (z)	-.014 (.030)	-.014 (.030)	-.014 (.030)	-.014 (.030)
Court size (z)	-.032* (.013)	-.032* (.013)	-.033* (.013)	-.032* (.013)
Random Effects	SD	SD	SD	SD
Level-2	.133***	.133***	.133***	.133***
Female	.165***	.151***	.148***	.142***
Level-1	.823***	.823***	.823***	.823***
N	134,602	134,602	134,602	134,602

⁺p<0.1, *p <0.05, **p<0.01, ***p<0.001

effect favoring female defendants; more specifically, female defendants on average received sentences that were 25.7 percent shorter than those imposed on male defendants. The effect size for gender reported here appeared to be significantly greater than that of prior federal sentencing studies, primarily because departure status was not controlled for in this model (see Ulmer et al. 2011b for instance). Interestingly, the magnitudes of the gender effect also differed substantially across federal district courts, as evidenced by the significant random variance component for gender ($SD=0.165$). The calculated 95% plausible range of the female effect indicated that 95% of districts had the gender effect, ranging from -0.580 to 0.066.¹⁷ Theoretically, this suggests the possibility that, in some districts, female defendants may be likely to be treated as severe as their male counterparts, while other female defendants, adjudicated in district situated on the other end of continuum, may enjoy an almost 60% shorter sentence relative to their male counterparts. The important question that this study attempts to answer, of course, is whether the two contextual variables, religious and political conservatism, could explain this variation.

Turning to the results on the interaction effects involving district-level religious (H-2) and political (H-3) conservatism, along with the complex three way interaction involving gender, religious and political conservatism (H-4), the results are reported in Model-2 through Model-4 in Table-6. To start with hypothesis-2, the results depicted in Model-2 provide strong support for the hypothesis-2. Consistent with expectations, there

¹⁷ The formula to calculate the range is $(-0.257 \pm (1.96*0.165))$.

was a significant interaction between gender and the level of religious conservatism, indicating that the effect of conservative religion reduced the effect of gender by 6.2 percent. The graphical representation of this interaction, presented in Figure-7, suggests that the gender discount, the gap in sentence severity between males and females, depends in part on the level of religious conservatism. That is, the gap was substantially larger in a district characterized by lower levels of religious conservatism. In the few districts with extremely high values of religious conservatism, however, the preferential treatment given toward female defendants was quite small. With regard to hypothesis-3, the results presented in Model-3 and Figure-7 revealed a quite similar pattern, as the political conservatism measure decreased the level of the female discount by 7.2%.

Figure-7 Cross-level interactions between gender and religious/political conservatism

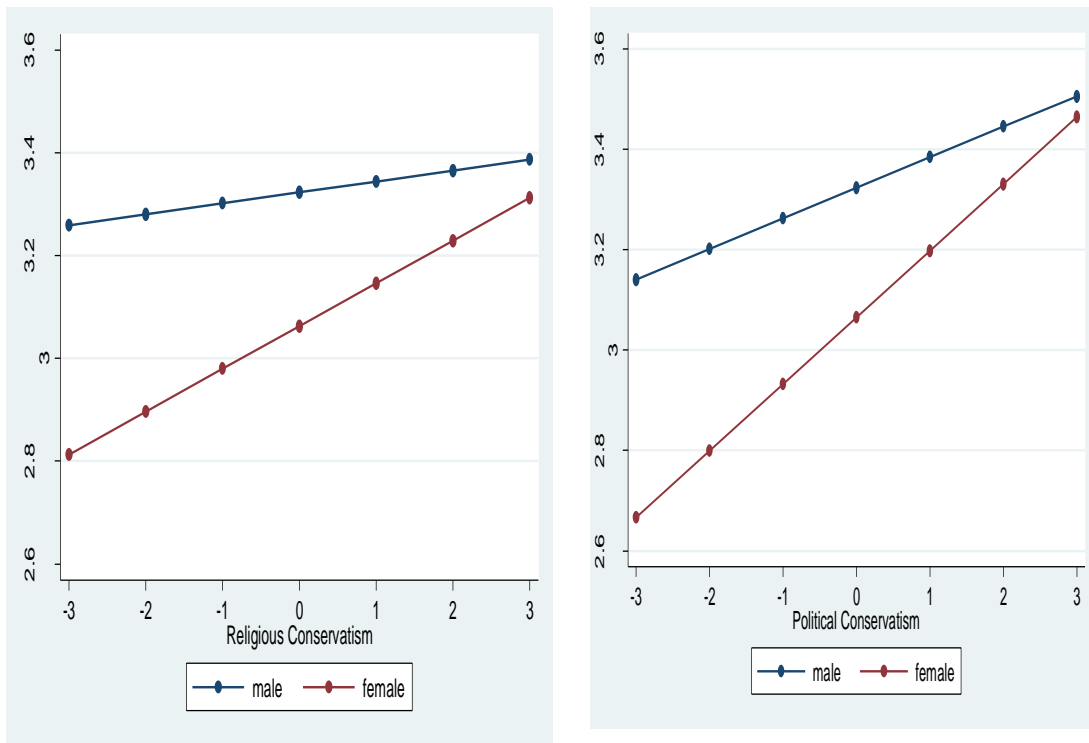


Figure-7 also clearly demonstrates that the gap in sentence length between men and women were conditioned by the level of political conservatism. For instance, the gender differential was quite minimal in a district characterized by an extremely high value of political conservatism. Finally, the finding described in Model-4 regarding the hypothesis on the three-way interaction (H-5) failed to receive empirical support, as there was no statistically significant interaction.

Supplemental Analyses

Before moving onto the results on the last two hypotheses of STUDY-1, I will provide supplementary analyses estimated using an alternative modeling scheme. As stated, given the recent methodological debate on the treatment of the departure status as a control variable, it is important to make a comparison between two models: one without departure controls and the other one with departure controls. The purpose of this part of analyses serves to check the robustness of the findings presented in Table-6 and, at the same time, to add to the recent discussion on the modeling issues concerning the roles of departure status in sentencing literature. The main findings for the model comparison are provided in Table-7, with each model having two columns representing different sub-models. To start, the results presented in Model-1 reveal a substantial divergence between the two regression coefficients for female: -0.257 from the model without a departure control v. -0.166 from the model with a departure control. The use of departure status as a control variable resulted in approximately a 35.4% reduction in the magnitude of the female effect. Similar patterns were also observed with the estimates for the two main

effects of the contextual variables. Overall, the effects of the two contextual variables on logged sentence lengths were significantly underestimated in the models with the departure control.

Table-7 Model Comparisons by Departure Controls (Abbreviated)

Fixed Effects	Model 1		Model 2		Model 3		Model4	
	NO	YES	NO	YES	NO	YES	NO	YES
Female	-.257***	-.166***	-.260***	-.169***	-.258***	-.167***	-.256***	-.165***
Religion	.023	-.003	.021	-.005	.023	-.003	.029	.004
Politics	.063***	.029*	.063***	.030*	.061***	.028*	.057***	.022
Female*Religion	-	-	.062***	.056***	-	-	.043*	.037*
Female*Politics	-	-	-	-	.072***	.071***	.049*	.051**
Religion*Politics	-	-	-	-	-	-	-.009	-.011
F*R*P	-	-	-	-	-	-	-.008	-.009

All the level 1 and level 2 variables are controlled, but not shown.

⁺p<0.1, *p <0.05, **p<0.01, ***p<0.001

Moving on to the results related to the interactions between gender and religious/political conservatism, unlike the marked differences in the main effects of gender and two other contextual variables, the results for the interactions presented in Model-2 and Model-3 turned out to be somewhat similar across two models. First, the two coefficients for the interaction between female and religious conservatism were similar in size and both of them were statistically significant (0.062*** v. 0.056***). Furthermore, the other two interaction terms between gender and political conservatism were similar and the magnitude of interaction effects were almost identical across the two

models (0.072*** v. 0.071***). Taken together, the overall findings presented in Table-7 provided further support for hypotheses-1, -2 and -3.

Race/Ethnicity-Specific Results: Hypothesis-5 and -6

As noted earlier, the final interest of the current study lies in uncovering race/ethnicity specific relationships, which may be masked in the pooled analyses. More specifically, this inquiry concerns the issue of to what extent the interaction effect of gender and the social contexts varies across race/ethnicity (H-5, H-6).¹⁸ In order to answer the two final questions, full sets of race/ethnicity-specific models were estimated and the results are provided in Table-8 and Table-9. To begin, Table-8 presents the results on cross-level interactions involving female and religious conservatism, with the first column indicating the result estimated from the pooled sample reported earlier. A review of the race/ethnicity-specific models suggested that the moderate effect of the cross-level interaction observed in the pooled sample were largely influenced by the effect for the African American sample; to elaborate, the effect of religious conservatism did not substantially reduce the preferential treatment given toward Hispanic (b=.049) and white (b=.051) female defendants, as the magnitudes of the interaction terms were smaller than that of the interaction term derived from the pooled sample (b=.062). In contrast, religious conservatism significantly decreased the negative effect of gender in the African American sample (b=.096). To provide more robust evidence of group

¹⁸ Given the small number of Asian, Native American, and other categories, I only limit the analysis to white, African American, and Hispanic populations to assess the final two hypotheses.

differences, a z-test was conducted to ascertain if the differences in the regression coefficients for the interaction terms between African Americans/Hispanic and white were statistically significant. The result showed that the gap in the effects of cross-level interactions indeed varied between the African American and white samples ($z=1.672$, $p=0.047$, one-tailed).

Table-8 Cross-level Interaction between Gender and Religion by Race/Ethnicity

Variables	Pooled sample	white	Black	Hispanic
Intercept	3.295*** (.028)	3.250*** (.038)	3.313*** (.032)	3.263*** (.035)
Female	-.260*** (.017)	-.218*** (.020)	-.327*** (.028)	-.302*** (.030)
Religion	.021 (.019)	.027 (.022)	.010 (.019)	.016 (.021)
Female*Religion	.062*** (.016)	.051** (.018)	.096*** (.020)	.049 ⁺ (.028)
N	134,802	48,289	42,029	37,831

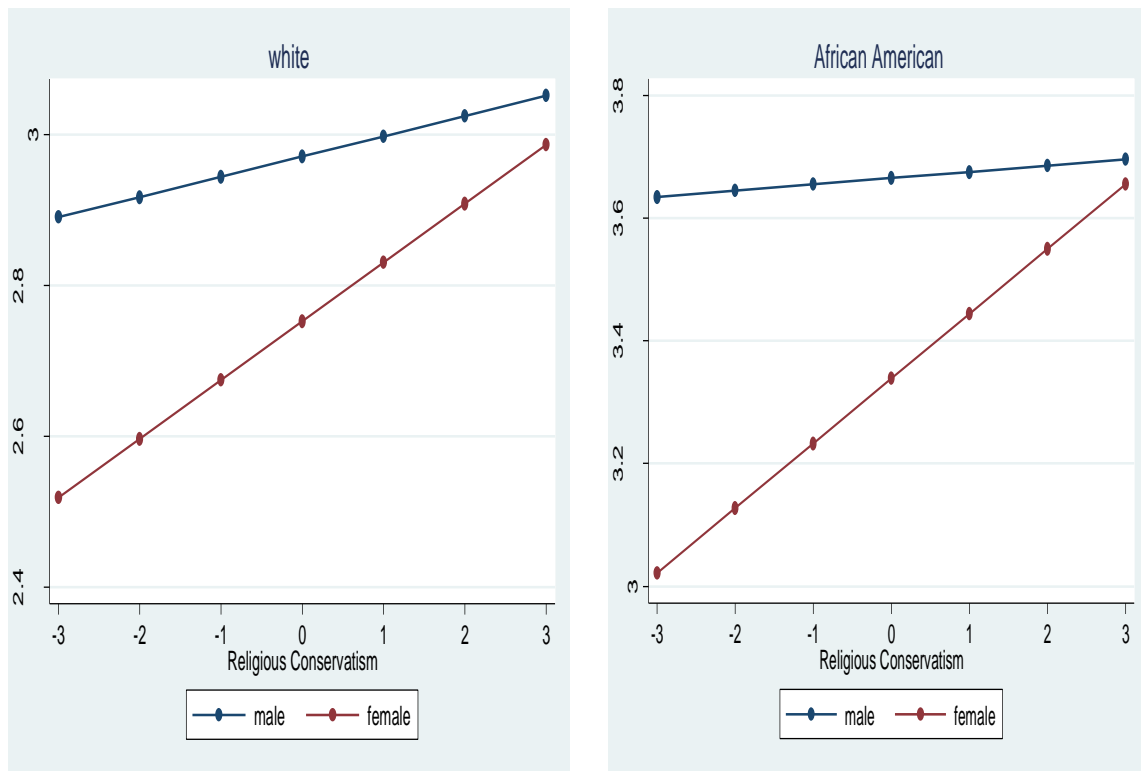
All the level 1 and level 2 variables are controlled, but not shown.

⁺ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

In order to help to interpret this complex relationship, Figure-8 is provided, which described the cross-level interactions between gender and religious conservatism for whites (left) and African Americans (right). It demonstrated that the degree to which the female discount depended on the level of religious conservatism significantly differed between Black and white offenders. A more careful look at the graph revealed a couple of interesting findings: first, the logged sentence lengths for male defendants, especially African American males, did not change meaningfully conditional on the level of religious conservatism, as the fitting line representing African American male was almost

flat. Second, the level of preferential treatment given toward female offenders, especially African American female offenders, was dramatically shaped by the religious context of the federal districts. The logged sentence length imposed on African American female offenders sentenced in a district characterized by lower levels of religious conservatism was substantially smaller than that imposed on African American female offenders sentenced in a district marked by higher levels of religious conservatism.

Figure-8 Cross-level interaction between gender and religious conservatism by race



Finally, turning to the results regarding gender and political context (H-6), the first column of Table-9 presents the findings from the pooled sample analyses provided earlier, indicating that there was a significant interaction effect between gender and political conservatism ($b=0.072$). More directly related to hypothesis-6, the

race/ethnicity-specific results presented in Table-9 and Figure-9 also revealed interesting patterns.

Table-9 Cross-level Interaction between Gender and Politics by Race/Ethnicity

Variables	Pooled sample	white	Black	Hispanic
Intercept	3.295*** (.028)	3.250*** (.038)	3.312*** (.032)	3.263*** (.035)
Female	-.258*** (.017)	-.218*** (.020)	-.311*** (.028)	-.304*** (.029)
Politics	.061*** (.017)	.065** (.020)	.058** (.018)	.066*** (.018)
Female*Politics	.072*** (.017)	.053** (.019)	.100*** (.021)	.088*** (.026)
N	134,802	48,289	42,029	37,831

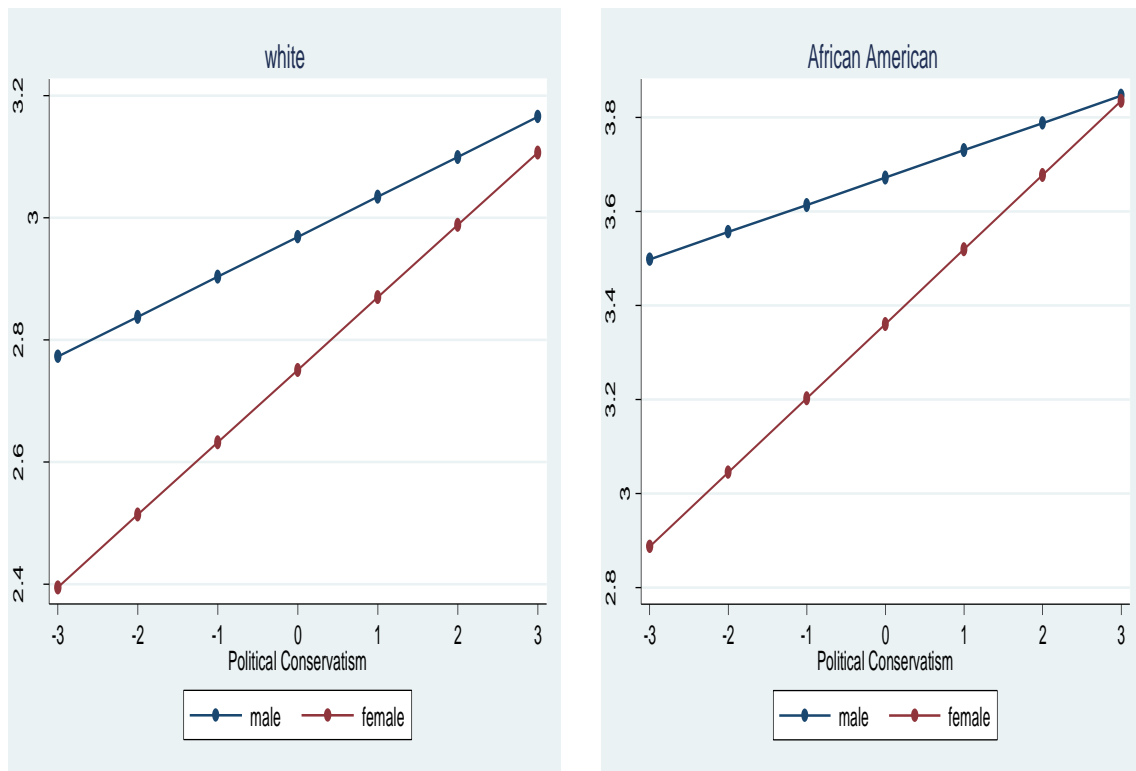
All the level 1 and level 2 variables are controlled, but not shown.

⁺p<0.1, *p<0.05, **p<0.01, ***p<0.001

First, it was discovered that the interaction term among African American defendants turned out to be greater than that of pooled sample (b=0.100); more specifically, Figure-9 suggested that the female discount that African American female defendants received was completely cancelled out in districts characterized by political conservatism three standard deviations above the mean. In contrast, African American females sentenced in a district court marked by political conservatism three standard deviations below the mean enjoyed the most substantial female discount. Second, it was also shown that the interaction between female and political conservatism among white offenders was also statistically significant, even though the magnitude of the interaction term was not greater than that of the interaction term estimated from the pooled sample (b=0.053). However, both male and female white defendants appeared similarly

influenced by the impact of the politically conservative context, as compared to their African American counterparts. In order to provide a more stringent test for the difference, a z-test was also performed. The results from the z-test also provided a support for hypothesis-6, as the difference between the regression coefficients for the interaction term was statistically significant at the p-value of 0.05 ($z=1.659$, $p=0.048$, one-tailed).

Figure-9 Cross-level interaction between gender and political conservatism by race



Summary of the Main Findings

To summarize, STUDY-1 found overall support for all the hypotheses proposed, except for hypothesis-4. More specifically, it was found that there existed a strong gender

effect favoring female defendants over male defendants at the sentence length decision (H-1). In addition, the effects of both religious and political conservatism significantly reduced the logged sentence length for female defendants (H-2 and H-3). However, the three-way interaction among the female, religion, and politics variables failed to receive empirical support (H-4). Finally, the race/ethnicity-specific analysis demonstrated that the cross-level interaction between gender and religious and political conservatism also significantly differed across whites and African Americans (H-5 and H-6). Overall, the results suggested that the sentences imposed on African American female offenders (relative to their male counterparts) were substantially conditioned by the two ecological contexts of court communities.

STUDY-2

Main Hypotheses

Hypothesis 1: The impact of gender will differ between the pre-Booker and the post-Booker periods such that the sentence discount that female defendants receive will be greater in the post-Booker period than in the pre-Booker period.

Hypothesis 2: The impact of gender will differ between the pre-Booker and the post-Gall periods such that the sentence discount that female defendants receive will be greater in the post-Gall period than in the pre-Booker period.

Hypothesis 3: Female defendants will enjoy less benefit of reduced sentence severity following Booker or Gall in a district court characterized by higher levels of religious conservatism.

Hypothesis 4: Female defendants will enjoy less benefit of reduced sentence severity following Booker or Gall in a district court characterized by higher levels of political conservatism.

Descriptive statistics

The descriptive statistics for all the variables employed in STUDY-2 were presented in Table-10. The dependent variable was the logged sentence length. Consistent with the approach taken earlier, this measure, which was capped at 470 months and log-transformed with a constant 1 added, had a mean of 3.09 with a standard deviation of 1.72. Turning to the independent variables, there are three key variables at each level that the current study employed. First, at level-1, female defendants accounted for approximately 15% of defendants. At level-2, another important independent variable concerned the Booker and Gall decisions. As noted earlier, this is a categorical variable with four levels of time indicator, representing Pre-Protect, Post-Protect, Booker and Gall. In the present sample, Booker cases accounted for the majority of the cases (31%), which was followed by Gall cases (29%), Pre-Protect cases (23%), and Post-Protect cases (17%). Finally, at level-3, there are two contextual level variables: religious and political conservatism, with the former having a mean of 188.17 and the latter having a mean of 57.

Control variables were employed at the individual (level-1), time (level-2), and district (level-3) levels. At level-1, most of the variables employed in STUDY-1 were also controlled for in the current study and descriptive statistics for these variables were largely comparable to those provided in Table-2. The only exceptions that warrant a brief

Table-10 Descriptive statistics

Dependent variable		Mean	SD	Min	Max
Sentence length (logged)		3.09	1.72	0	6.15
Independent and Control variables		Mean	SD	Min	Max
Level 1 Variables	Female	.15	.36	0	1
	White (reference group)	.36	.48	0	1
	African American	.30	.45	0	1
	Hispanic	.28	.45	0	1
	Others	.04	.21	0	1
	Dependent	.61	.48	0	1
	Age	35.16	11.29	16	102
	Some college	.25	.43	0	1
	Non-citizen	.20	.40	0	1
	Plea	.94	.21	0	1
	Detained	.65	.47	0	1
	Violent (reference group)	.04	.21	0	1
	Drug	.47	.49	0	1
	Fraud	.19	.39	0	1
	Firearms	.14	.34	0	1
	Property	.03	.18	0	1
	Others	.14		0	1
	Multiple Counts	.22	.41	0	1
	Criminal History	.72	.44	0	1
	Presumptive sentence (logged, GLMIN)	3.49	1.49	0	6.15
Presumptive sentence (logged, XMINSOR)	3.45	1.46	0	6.15	
Level 2 variables	Pre-Protect (reference group)	.23	.42	0	1
	Post-Protect	.17	.37	0	1
	Booker	.31	.46	0	1
	Gall	.29	.45	0	1
	% Republican Appointees	60.90	16.88	0	100
	% Male judges	79.06	12.09	50	100
	% white judges	80.90	14.55	47.05	100
	Caseload	3.96	3.30	.27	17.7
	Crime rates	25.19	13.79	.76	83.18
	Time	59.00	33.69	0	116
Level 3 variables	Conservative Religion	188.17	132.71	14.01	517.28
	Political Conservatism	57.00	8.72	34.55	77.45
	% Black residents	9.21	10.16	0.14	44.83
	Disadvantage (z)	0	1	-1.97	2.33
	Court size	10.55	6.62	1	28

All the variables are presented as their original forms, unless otherwise specified.

explanation were the two presumptive sentence variables shown in Table-10. As noted earlier, the first presumptive sentence variable was based on the GLMIN variable, which took into account the mandatory minimum sentence, which would trump the statutory minimum sentence. This variable had a mean of 3.49 and a standard deviation of 1.49. The second presumptive sentence variable was derived from the XMINSOR variable, which is not constrained by the mandatory minimum sentence and it had a mean of 3.45 and a standard deviation of 1.46.

At level-2 and level-3, the same district level variables used in Study-1 were employed. The only difference lies in the fact that some level-2 variables from Study-1 were used as time-varying covariates in the present study, thus effectively controlling for any variations which would change over time. For instance, the federal dockets were described as being presided over by mostly white and male judges. Crime rates had a mean of 25.19 with a standard deviation of 13.79. The descriptive statistics for all the level-3 variables were the same as those already explained previously in the discussion of STUDY-1.¹⁹

Unconditional Model

Before estimating full models to test each hypothesis proposed, I will start with an unconditional model. An unconditional three-level model allows for partitioning the total variability in the outcome into three distinct components and the significant variations

¹⁹ I am not presenting correlation matrixes for STUDY-2, as the estimates are quite similar to those reported in STUDY-1.

provide justification for estimating multilevel models. The results, presented in Table-11, indicated that statistically significant variations indeed existed in the average sentence length both at the district (level-3) and time (level-2) levels. To be more specific, the estimate of the random variance for districts showed that there was a statically significant random variation in sentence lengths across districts ($SD=0.382$, $p<.001$), which corresponded to an ICC of 0.049. This means that about 5% of the variance in logged sentence length could be attributed to the district in which the case was adjudicated. The random variance for time was also significant, even though the ICC was rather small ($SD=0.246$, $p<.001$). The small random variance for time might suggest that time-level clustering is trivial such that a complex three-level HLM is not necessary. In order to take into account this possibility, a likelihood ratio test was performed, which produced results in support of the three-level model over a simple two-level model in which cases were only nested within districts (LR chi-square= 4530.37 , $p<.001$).

Table-11 Unconditional Model

Fixed Effects	Coefficients	SE	P
Intercept	3.138***	.040	.000
Random Effects	SD	SE	P
Level-3	.382***	.028	.000
Level-2	.246***	.003	.000
Residual	1.659	.001	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Results regarding Hypothesis-1 and -2

The first two hypotheses ask whether there are changes in the effects of gender in the wake of the Booker and Gall decisions, as compared to the effect of this variable in the Pre-Protect period. Before presenting the answers to those questions directly, I will provide a baseline model to contextualize the findings related to hypothesis-1 and-2. First, review of Model-1 in Table-12 suggests that female defendants received a 24% shorter sentence relative to their male counterparts. To be more specific, as all the individual level variables, except for gender, were grand-mean centered and all the district level covariates were standardized, except for Booker and Gall, the coefficient for female represents the difference in the logged sentence lengths between average males and females sentenced in a district characterized by average characteristics during the Pre-Protect period. A majority of other findings related to the impacts of other covariates were largely consistent with the results reported in STUDY-1. At the district level, the estimates for the two main variables of interest also suggested that both religious and political conservatism were positively associated with the logged sentence length. Finally, the results depicted in Model-1 demonstrate the main effects of the Booker and Gall decision—that is, the Booker and Gall decisions led to reductions in logged sentence lengths by 6.7% and 8.1% respectively, as compared to the Pre-Protect period. These findings are largely consistent with some of the results reported in prior studies, which demonstrated a decline in sentence severity in the wake of Booker and Gall (see Fischman & Schanzenbach, 2012; Kim et al., Forthcoming; Rehavi & Starr, 2013). The

Table-12 Full Models

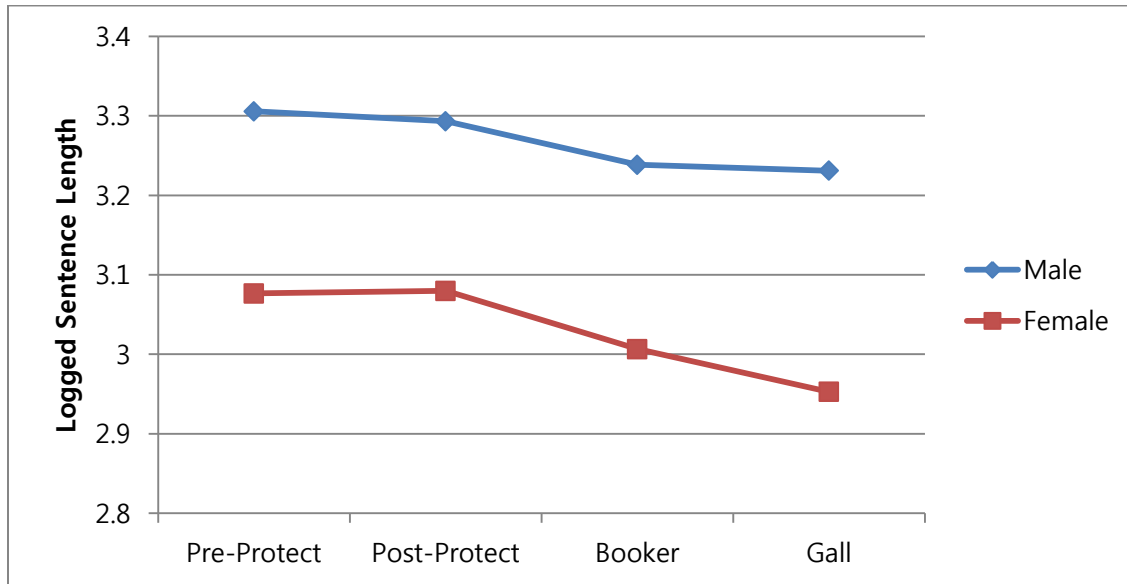
Fixed effects	Model-1			Model-2		
Female	-.241***	.003	.000	-.229***	.006	.000
Protect	-.010	.008	.241	-.012	.008	.241
Booker	-.067***	.012	.000	-.067***	.012	.000
Gall	-.081***	.016	.000	-.074***	.016	.000
Female*Protect	-	-	-	.015	.010	.135
Female*Booker	-	-	-	-.002	.009	.751
Female*Gall	-	-	-	-.048***	.009	.000
Intercept	3.233***	.018	.000	3.231***	.018	.000
Black	.103***	.003	.000	.103***	.003	.000
Hispanic	.064***	.003	.000	.064***	.003	.000
Others	.034***	.006	.000	.034***	.006	.000
Dependent	-.016***	.002	.000	-.016***	.002	.000
Age	-.002***	.000	.000	-.002***	.000	.000
Education	-.060***	.002	.000	-.060***	.002	.000
Non-citizen	.086***	.003	.000	.086***	.003	.000
Plea	-.294***	.005	.000	-.294***	.005	.000
Detained	.605***	.003	.000	.605***	.003	.000
Drug	-.121***	.005	.000	-.121***	.005	.000
Fraud	-.280***	.006	.000	-.280***	.006	.000
Firearms	-.062***	.006	.000	-.063***	.006	.000
Property	-.269***	.008	.000	-.270***	.008	.000
Others	-.187***	.007	.000	-.188***	.007	.000
Multiple Counts	.262***	.003	.000	.262***	.003	.000
Criminal history	.169***	.003	.000	.169***	.003	.000
Presumptive sentence	.827***	.001	.000	.827***	.001	.000
% Republican Judges (z)	-.005	.011	.656	-.005	.011	.656
% Male Judges (z)	.004	.013	.720	.005	.013	.720
% white judges (z)	.019	.017	.255	.019	.017	.255
Caseload (z)	.015*	.006	.011	.015*	.006	.011
Crime rates (z)	-.005	.003	.100	-.005	.003	.100
Time	.004***	.000	.000	.004***	.000	.000
Time ²	-.000***	.000	.000	-.000***	.000	.000
% Black Residents (z)	.023	.016	.150	.023	.016	.150
Disadvantage (z)	.003	.019	.848	.003	.019	.848
Court Size (z)	-.013	.012	.261	-.013	.012	.261
Religion (z)	.039*	.018	.034	.039*	.018	.034
Politics (z)	.061***	.017	.000	.061***	.017	.000
Random effects	SD	SE	P	SD	SE	P
Level-3	.131***	.010	.000	.132***	.010	.000
Level-2	.088***	.001	.000	.088***	.001	.000
Residual	.809***	.000	.000	.809***	.000	.000
N	463,883			463,883		

⁺p<0.10, *p <0.05, **p<0.01, ***p<0.001

important question, of course, is whether the decline was equal for male and female defendants, which I will discuss below.

Turing to the results concerning hypothesis-1 and -2, Model-2 in Table-12 indicates that there was a significant interaction effect between gender and Gall, thus confirming hypothesis-2. That is, the negative effect of being female ($b=-0.229$) became even greater during the Gall period by -0.048 , as compared to one during the Pre-Protect. In order to contextualize this finding, Figure-10 is provided, which graphically demonstrates how sentence severity for males and females changed over time especially under the contexts of Booker and Gall. To start, it was clear that there was a reduction in sentence severity in the wake of the Booker and Gall decisions, as there was a downward trend in the adjusted means for the logged sentence length for Booker and Gall. Turing directly to the findings related to hypothesis-1 and -2, the effect of gender—the gap in the logged sentence length between males and females—appeared similar during the Booker period, as compared to the gap during the Pre-Protect period. In other words, even though there was an overall reduction in sentence severity following the Booker decision, this decrease was largely comparable for males and females, thus failing to provide support for hypothesis-1. However, Figure-10 suggests that the average female discount increased during the Gall period, as compared to the Pre-Protect period. In the wake of the Gall decision, the average sentence lengths imposed on male defendants did not change, while the average sentences imposed on female defendants kept declining. Overall the results depicted in Table-12, along with Figure-10 provide support for hypothesis-2.

Figure-10 Interaction between gender and Booker/Gall



Supplemental Analyses

As noted earlier, one of the important methodological debates in the Booker/Gall literature is the treatment of mandatory minimum status. Some scholars criticize the prior studies for employing mandatory minimum status, along with departure status, as control variables due to a potential endogeneity issue (see Fischman & Schanzenbach, 2012; Starr & Rehavi, 2013). Against this backdrop, in this section I provide results from two different models with the main difference being the measurement of the presumptive sentence. The results provided in the left column of each model in Table-13 are derived from a model in which the presumptive sentence based on XMINSOR was employed, thus not controlling for mandatory minimum status. By contrast, the results presented in the right column of each model utilized a measure of the presumptive sentence based on

GLMIN, therefore effectively making the application of mandatory minimums constant across different periods.

Table-13 Model Comparisons by Mandatory Minimum Controls

Fixed Effects	Model 1		Model 2	
	NO	YES	NO	YES
Female	-.241***	-.235***	-.229***	-.224***
Protect	-.010	-.008	-.012	-.012
Booker	-.067***	-.068***	-.067***	-.068***
Gall	-.081***	-.093***	-.074***	-.086***
Female*Protect	-	-	.015	.023*
Female*Booker	-	-	-.002	-.001
Female*Gall	-	-	-.048***	-.051***

All the level-1, level-2, and level-3 variables are controlled, but not shown

+p<0.1, *p <0.05, **p<0.01, ***p<0.001

To start with the findings in Model-1, the results suggest that the model controlling for mandatory minimums slightly over-estimated the impacts of Booker and especially Gall, because this model did not take into account any changes in the application of mandatory minimums. This finding is consistent with the prior studies reporting that federal prosecutors were increasingly relying on the use of mandatory minimums to counter the impact of increased judicial discretion in the wake of Gall (see Fischman & Schanzenbach, 2012). A similar pattern is also observed in the results depicted in Model-2. The results suggest that the model incorporating mandatory minimum status slightly over-estimated the interaction between gender and Gall. Taken

together, the overall results provided in Table-13 lend further support to hypothesis-2, as the finding was robust to a different model specification. In addition, the finding also suggests that the use of a measure of the presumptive sentence that incorporates the mandatory minimum sentence especially in assessing the impacts of Booker/Gall would lead to a biased estimate of the variables of interest.

Results regarding Hypothesis-3 and -4

The last two research questions focus on whether the magnitude of the interaction effect between gender and Booker/Gall depends on the level of religious/political conservatism in the district in which the case was adjudicated. Connected with the findings reported earlier regarding hypothesis-1 and-2, this means that the extent to which the gender gap in logged sentence length increases in the wake of the Gall decision would be systemically different across levels of religious/political conservatism. Hypothesis-3 and-4 predicted that the negative interaction effect between gender and Gall would be significantly smaller in districts characterized by higher levels of religious and political conservatism.

Given the complexity of the hypothesized relationships, I will provide a table along with two different types of graphs to help dissect the three-way interactions. Model-1 in Table-14 depicts the results regarding hypothesis-3. As the previous analyses provided support for the hypothesis on the existence on the two-way interaction between gender and Gall, but not Booker, the main interest of the hypothesis-3 concerns whether a three-way interaction among gender, Gall, and religious conservatism is statistically

significant. The result show that there was indeed a statistically significant three-way interaction (b=0.024, p=.007). A more robust way to check the existence of a three-way interaction is to examine the significance of an overall interaction, not just to look at the specific interaction term, because the policy change is a categorical variable with four levels. The contrast command provided by STATA-13 also confirmed the existence of an overall interaction among gender, policy changes and religious conservatism evidenced by a significant chi-square test (chi-square=10.71, df=3, p=.013).

Table-14 Three-way Cross-level Interactions

Model-1 (Religion)			Model-2 (Politics)		
Fixed effects	Coefficient	SE	Fixed effects	Coefficient	SE
Female	-.228***	.006	Female	-.227***	.006
Protect	-.011	.008	Protect	-.011	.008
Booker	-.066***	.012	Booker	-.065***	.012
Gall	-.073***	.016	Gall	-.073***	.016
Religion	.021	.018	Politics	.037*	.017
Female*Protect	.015	.010	Female*Protect	.013	.010
Female*Booker	-.004	.009	Female*Booker	-.009	.009
Female*Gall	-.054***	.009	Female*Gall	-.056***	.009
Female*Religion	.036***	.006	Female*Politics	.052***	.006
Protect*Religion	.005	.005	Protect* Politics	.010	.005
Booker*Religion	.010*	.004	Booker*Politics	.021***	.004
Gall*Religion	.017***	.004	Gall*Politics	.020***	.004
Female*Protect*Religion	-.003	.010	Female*Protect*Politics	.008	.010
Female*Booker*Religion	.007	.008	Female*Booker*Politics	.025**	.008
Female*Gall*Religion	.024**	.009	Female*Gall*Politics	.020*	.009
Intercept	3.229	.018	Intercept	3.228	.018
Random effects	SD	SE	Random effects	SD	SE
Level-3	.131***	.010	Level-3	.131***	.010
Level-2	.088***	.001	Level-2	.088***	.001
Residual	.808***	.000	Residual	.808***	.000

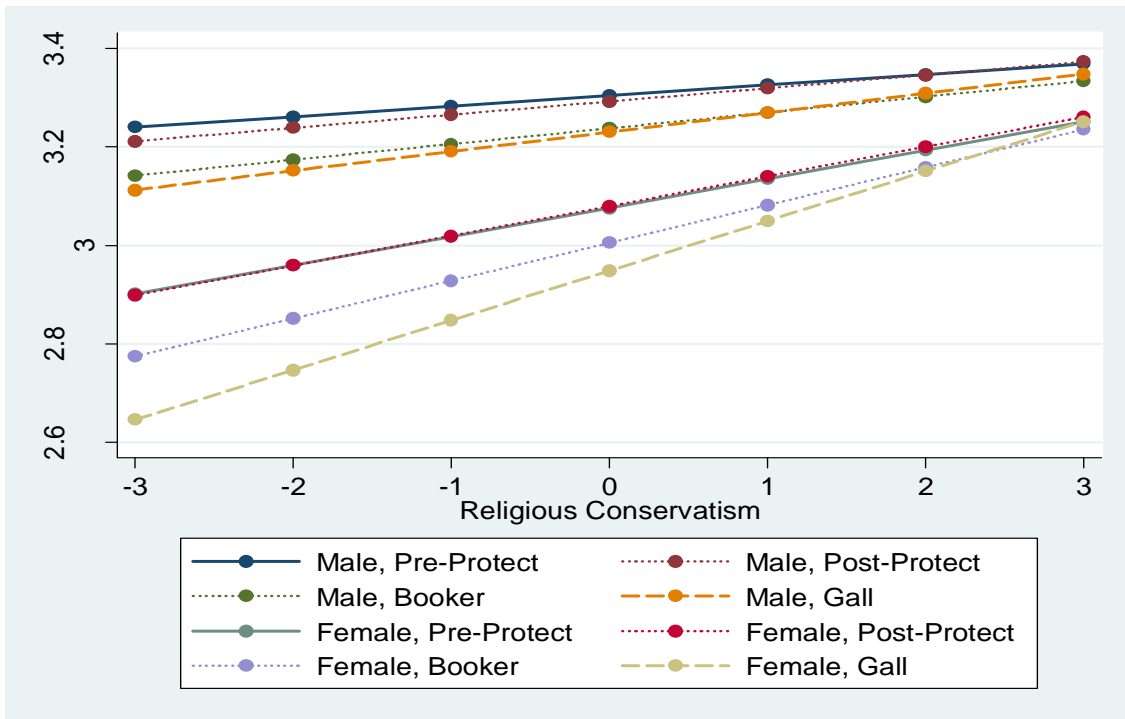
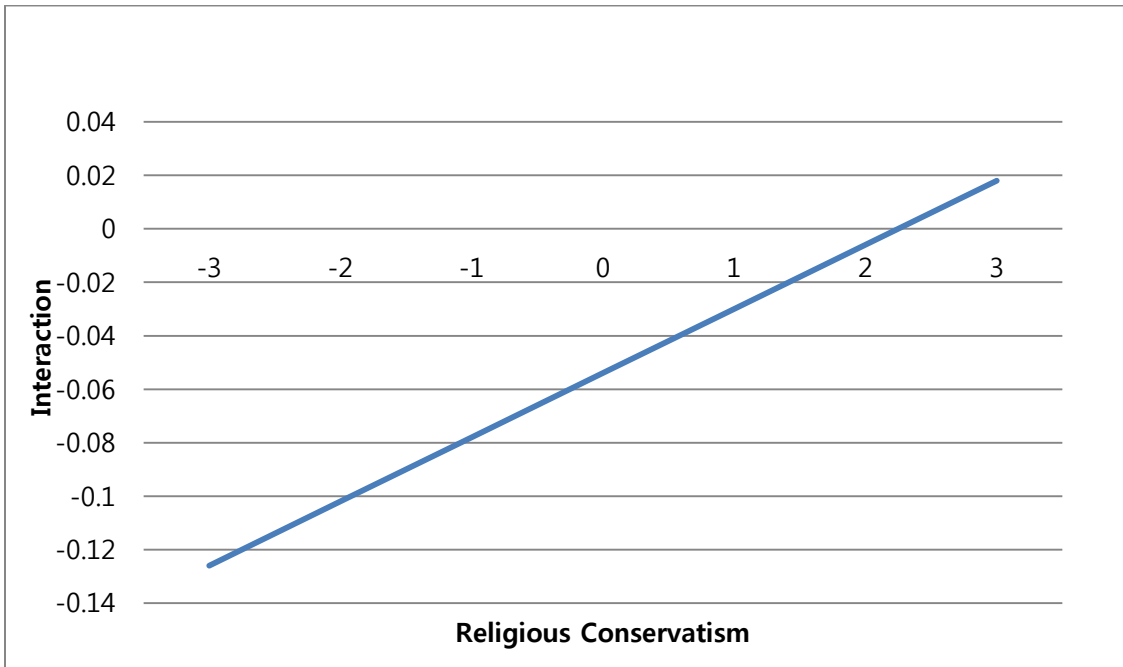
All the level-1, level-2, and level-3 variables are controlled, but not shown.

⁺ p <0.1, *p <0.05, **p<0.01, ***p<0.001

As noted, I present two graphs to better illustrate the three-way interaction. The upper graph in Figure-1 shows that the magnitude of the interaction (shown on the y-axis) depends on the level of religious conservatism (shown on the x-axis). To be more specific, the negative interaction between gender and Gall—that is, a reduction in sentence length following the Gall decision for female offenders—was even greater in federal districts characterized by lower levels of religious conservatism. By contrast, the negative interaction disappeared in districts characterized by higher levels of religious conservatism. If there had been no three-way interaction among these variables, the slope for religious conservatism would have been flat. A more straight-forward result can be found in the lower graph in Figure-11, where the two solid lines represent males and females sentenced in the Pre-protect period (reference category) and the two dash lines represent males and females sentenced in the Gall period.²⁰ This graph demonstrates that the gender gap in the logged sentence length between males and females during the Gall period was significantly greater in districts characterized by a lower level of religious conservatism, as compared to the gender gap during the Pre-Protect period. In contrast, the female discount found in the Gall period was not dramatically different from the one found in the Pre-Protect period among the defendants sentenced in districts with higher levels of religious conservatism.

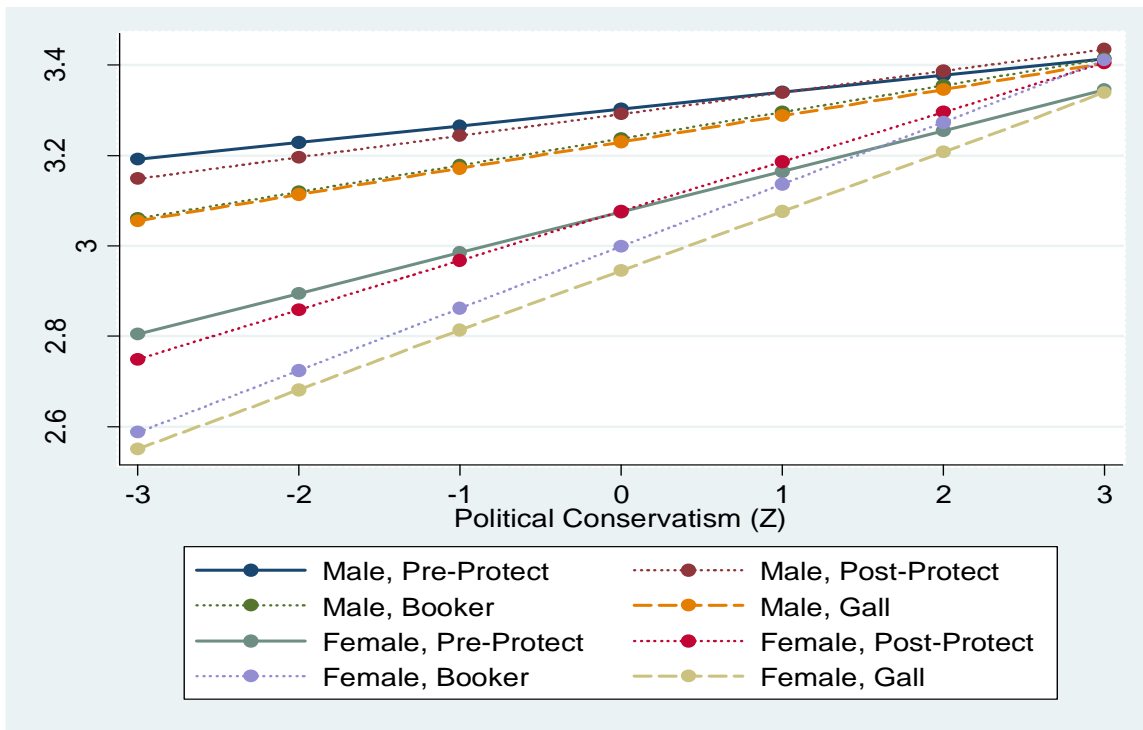
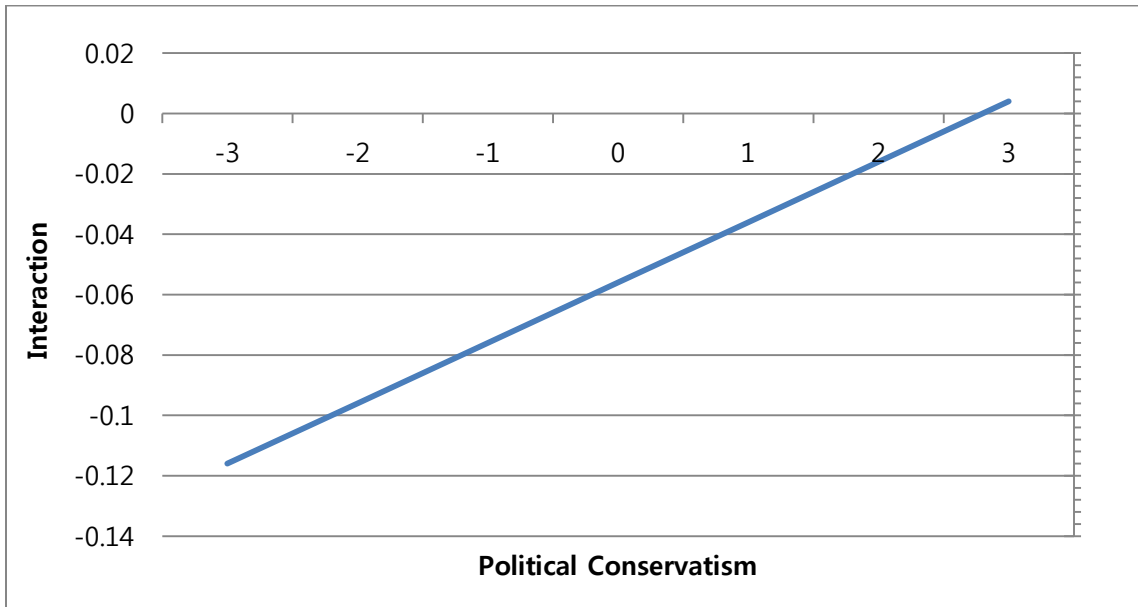
²⁰ It should be noted that the y-axis for this graph is the logged sentence length, while the y-axis for the upper graph is the interaction between gender and Gall.

Figure-11 Three-way interaction-religious conservatism, gender, Booker/Gall



The final research question involves a three-way interaction among gender, Booker/Gall, and political conservatism. The results provided in Model-2 in Table-14 were largely consistent with the findings presented earlier in Model-1. Findings lend support to hypothesis-4, as there is a statistically significant three-way interaction among gender, Gall, and political conservatism ($b=0.020$, $p=0.026$). In a similar way to the three-way interaction among gender, Gall, and religious conservatism, the effect of the negative interaction between gender and Gall—the degree to which the gender gap increased in the wake of Gall—was mitigated by the political conservatism of the district. One prominent exception to this finding of a similar pattern across Model-1 and Model-2 is that the three-way interaction among gender, Booker, and political conservatism was significant ($b=0.025$, $p=0.003$), despite the absence of the two-way interaction between gender and Booker. This suggests that the interaction effect involving Booker and gender, which was not statically significant at the p -value of .05, was highly conditional on the effect of political conservatism. Finally, the contrast command also reconfirmed the existence of an overall interaction among gender, policy changes and political conservatism, as evidenced by a significant chi-square test (chi-square=9.45, $df=3$, $p=0.023$). The two graphs included in Figure-12 also clearly demonstrate the relationship among three variables.

Figure-12 Three-way interaction-political conservatism, gender, Booker/Gall



Summary of the Main Findings

To summarize, STUDY-2 found a significant interaction between gender and Gall, but not gender and Booker. These findings provide support for hypothesis-2. Taken together, the results regarding hypothesis-1 and -2 reveal that female defendants were the recipients of even greater leniency during the Gall period, but not the Booker period, relative to the Pre-Protect period. With regard to the latter two hypotheses, this study found partial support for hypothesis-3. There was a statistically significant three-way interaction among gender, Gall, and religious conservatism. That is, the additional female discount that female defendants enjoyed during the Gall period was substantially greater in districts characterized by lower levels of religious conservatism. I also found convincing support for hypothesis-4, as there were two statistically significant three-way interactions involving both Booker and Gall. The three-way interaction among gender, Booker, and political conservatism was particularly noteworthy, because there was no two-way interaction between gender and Booker.

CHAPTER 6: DISCUSSION AND CONCLUSION

Discussion of the Main Findings

Literature on gender disparity in sentencing has a long history, raising a seemingly simple but ultimately complex question, such as “should women and men be treated equally in courtroom?” (Roberts, 1994). The focus of this dissertation, however, is to move beyond this traditional micro approach by exploring broader social contexts which shape gender disparity in sentencing outcomes. More specifically, I examine how religious and political conservatism condition the impact of gender on sentencing outcomes in federal district courts. Below, I discuss the implications of the main findings of this dissertation, along with the strengths and weaknesses of this study, which will be followed by theoretical and policy implications flowing from this dissertation. Finally, I will finish this dissertation with the conclusion.

STUDY-1

The purpose of STUDY-1 was to move beyond prior research on gender and sentencing by examining the roles of two important elements of social contexts—religious and political conservatism—that potentially shape gender disparity in sentencing outcomes. In doing so, more specific attention was devoted to how the relationship between the conservative social climates and gender in courtroom decision-making is further

contextualized by the race and ethnicity of offenders. Even though a lot could be discussed with regard to the main findings of STUDY-1, some of the most important findings which warrant further attention are as follows.

To begin, a basic, but noteworthy, finding revealed in this dissertation is that there is a strong gender effect favoring female over male offenders. On average, the sentences imposed on women offenders were about 26% shorter than those imposed on male offenders. As mentioned, the magnitude of the gender effect is substantively greater than the gender effect reported in previous studies in federal sentencing literature (see for instance Ulmer et al., 2012b). This difference is mostly attributable to a methodological choice which involves the inclusion of a variable controlling for whether the offender received a downward departure. In line with the serious criticism raised by a group of legal scholars (see Starr & Rehavi, 2014), this study found that the use of the departure control variable indeed influenced the size of the extra-legal disparity of interest in a meaningful manner. I therefore call for a serious discussion about some of the methodological issues that sentencing literature faces to advance the literature to the next level.

What is more relevant with regard to the main topic of this dissertation is the substantial variability of the gender effect across federal district courts. This finding provided a unique opportunity to dissect the gender disparity through a contextual lens that has not been utilized before (see Bontrager, 2013). Indeed, this dissertation found that the gender differential in sentence severity was significantly smaller in court

communities characterized by higher levels of religious and political conservatism. Generally speaking, these variables had similar conditioning effects on the gender differential in sentence severity and I interpreted the findings as providing support for the main predictions from the focal concerns and court community perspectives.

This study also revealed that, as consistent with the theoretical expectations on the relationship between religious and political conservatism (see King, 2008), there were some differences involving the roles of the two social contextual variables in judicial decision-making. This study found no main effect of religious conservatism. Even though investigating the main effect was not a major concern for this dissertation, this null finding does not neatly align with the theoretical literature arguing for a close relationship between Christian fundamentalism and punishment severity (see Grasmick et al., 1992; Myers, 1989). However, as consistent with the theoretical expectation, this study found that the effect of religious conservatism was conditional on the gender of offenders. What is intriguing is the pattern of the cross-level interaction involving gender and religious conservatism; religious conservatism affected the punishment levels for female, but not for male offenders. These findings provide a reason to believe that the effect of religious conservatism is largely mediated by gender norms (see Moore & Vanneman, 2003), which in turn determine the level of social control directed at female offenders, but not male offenders.

A more important question, of course, is how conservative gender attitudes shape judicial decision-making, as federal judges are trained as professional legal practitioners

not to invoke societal stereotypes. Although I can only speculate, there are a couple of possibilities that may be at work, either independently or simultaneously. The first possibility concerns the use of implicit gender bias. As briefly reviewed earlier, recent work in social psychology uncovered that, in many cases, people do not overtly rely on social stereotypes, but rather the use of social stereotypes is implicit (Greenwald et al., 1998).

Extending this recent body of work, it may be the case that implicit bias plays a larger role, especially when coupled with social environments more favorable to the use of implicit bias. Thus, it is plausible that judges from courts embedded in communities with higher levels of religious conservatism are more likely to rely on gender stereotypes in their decision-making, as compared to the judges from less conservative court communities. Another potential and more traditional approach would be the possibility that local gender norms prevalent in court environments are built into local legal culture. That is, substantive rationality built into the local legal culture dictated specific going rates for women offenders. Thus, judges in the federal districts characterized by higher levels of religious conservatism may have been more likely to mete out more severe sentences to female offenders, as compared to female offenders adjudicated in less conservative court communities.

Turning to the findings related to political conservatism, it was interesting that, unlike religious conservatism, this study found that political conservatism had both main and interaction effects on sentencing outcomes. That is, political conservatism was found

to increase the average level of sentence severity and to reduce the female discount, adding to the growing body of literature demonstrating the efficacy of political conservatism in sentencing literature (Garland, 2002; Jacobs & Helms, 1996; Mauer, 2001). Upon closer examination, it is worthwhile to point out that political conservatism affected sentence outcomes for both male and female offenders, albeit to a different degree. This makes an interesting comparison with regard to the impact of religious conservatism for male and female offenders. It is possible that the impact of political conservatism may have worked through the concept of individual responsibility inculcated in conservative political values (King, 2008). As reviewed earlier, political conservatism underscores the importance of individual responsibility in such things as crime and poverty (Garland, 2001). Thus, in a court community characterized by a higher level of political conservatism, individual responsibility may be heavily weighted by judges, with all other factors closely related to the unique circumstances of females receiving little or no attention. By contrast, in court communities marked by a higher level of political liberalism, it is plausible that court actors recognize that men and women live in worlds characterized by different social realities (Chesney-Lind & Pollock, 1995); as a result, judges may consider the social adversities surrounding women in their sentencing decisions, such as being a single mom etc.

This reasoning is in line with the prior literature (see King, 2008), documenting that the underlying factors driving the impact of political conservatism on formal social control may be two-fold: moral values (Garland, 2001) and local partisan strategies (Beckett, 1997). Sentencing literature also suggests that a primary channel by which local

political conservatism affects sentencing outcomes is the local political processes that influence the selection of judges and prosecutors (Eisenstein et al., 1988; Eisenstein & Jacob, 1977). In other words, court communities with higher levels of political conservatism are more likely to have judges and other courtroom workgroup members with conservative political affiliations. Because this dissertation controlled for a proxy variable measuring the political affiliations of the federal judges, along with the gender and race of the judges, the possibility that the impact of local political conservatism works through the local partisan strategy could be effectively ruled out. Thus, it is probable that moral beliefs emphasizing individual responsibility enmeshed in political conservatism affect judicial decision-making for both male and female offenders.

Another interesting finding is the fact that the influence of social contexts on judicial decision-making has race-specific implications. It was found that the treatment of African American female offenders was significantly more likely to be contextualized by both religious and political conservatism, as compared to the treatment for white female offenders. Of particular relevance is the finding related to race and religious conservatism. That is, regardless of the levels of religious conservatism, the sentence lengths imposed on African American males did not differ significantly, while the treatment of African American females was significantly harsher in court communities with higher levels of religious conservatism.

There may be two inter-related factors contributing to this intriguing pattern. To start with the finding related to the context-free effect concerning African American

males, the race and sentencing literature has long described (young) African American males as a potential social threat (Liska, 1992) and many studies also demonstrated that this group of offenders was more likely to be singled out for harsher punishments (see Steffensmeier et al., 1998; Spohn & Holleran, 2000). Regarding the finding on the relationship between religious conservatism and African American female offenders, there is an opposing approach which views African American females as in need of social help. Indeed, many African American religious groups call for a more egalitarian social approach and collective action that emphasize the different roles of African American females (Wilcox & Thomas, 1992). Once again, this finding provides further support for the speculation that religious contexts have more gender specific implications, as compared to political contexts.

Taken together, these findings on the interconnectedness among gender, race, and social contexts suggests that the traditional topic of whether chivalry bypasses women of color needs to be expanded. That is, the relationship between gender and race should be reassessed through a contextual lens, as the connection is more likely to be interpreted distinctively under different social contexts. In that sense, this finding is quite consistent with the argument that the concepts of race and ethnicity are quite *fluid* (Ulmer, 2012).

STUDY-2

The second study sought to build on the first study by examining the influence of social contexts against the backdrop of policy changes. As discussed before, one of the

key hypotheses derived from the court community perspective, which has been empirically neglected over the years, is that the impact of sentencing reforms or policy changes is almost always filtered through the local norms of court communities. In that regard, the second study attempts to address this limitation (see Kim et al., forthcoming; Reiter & Frank, 2013 for exceptions).

Although the key empirical foundation of STUDY-1 was built on the main effect of gender and its variation across federal districts, the starting point for STUDY-2 was how the gender disparity in sentence severity might change across landmark decisions such as Booker and Gall. To begin, this study revealed that there was a significant reduction in sentence severity following Booker and Gall and, more importantly, the decline was more pronounced for female offenders during the Gall period, but not the Booker period, relative to the Pre-Protect period. Meanwhile, in order to situate the findings within a broader context, it is necessary to revisit the study by the USSC (2004) and a small group of studies conducted in other states assessing the impact of policy changes on gender disparity (see Griffin & Wooldredge, 2006; Koons-Witt, 2002; Miethe & Moore, 1985). As reviewed earlier, the USSC (2004) study indicated that gender disparity appeared to increase following the implementation of the federal sentencing guidelines, a result that clearly conflicts with the conclusions of other studies reporting that implementation of guidelines had a limited effect on gender disparity in sentencing.

The study by Griffin and Wooldredge (2006) provides a valuable insight on why gender disparity continues to grow in federal district courts following the federal

guidelines' transition into advisory guidelines. Griffin and Wooldredge attribute the relative stability of gender effects across the sentencing reform in Ohio to wide discretion allowed to judges and to the overall sentence severity built into the guidelines, which reflected the pre-guidelines sentencing patterns. In contrast, in federal district courts characterized by the guidelines' excessive rigidity and overly harsh punishments, judges may have felt that they were not allowed to impose sentences that fit more with the culpability of female offenders. In a similar vein, Kramer and Ulmer (2009) assert that when there is a mismatch between local actors' definitions of offender blameworthiness and dangerousness, judges view the sentencing guidelines as simply too severe for such cases, and do not follow the guidelines. In addition, the implication of the Gall decision is also nicely tied with the federal guidelines' rigid stance on gender and gender-related factors. Unlike its predecessor, Booker, Gall made it clear that policy disagreement could be a legitimate reason for guidelines departures. Therefore, it is quite possible that a majority of judges do not consider the prohibition against using gender-related factors warranted and therefore depart from the guidelines when sentencing female offenders (see Raeder, 2006).

Turning to the main topic of Study 2, this study found that, consistent with theoretical expectations, the policy changes were modified by the social contexts of federal court communities. More specifically, there were larger increases in gender disparity following the Gall decision in districts with lower levels of religious and political conservatism. Another interesting finding emerged with regard to the conditional effect of political conservatism during the Booker period. It was found that there was no

significant interaction between Booker and gender, but it turned out that there was a statistically significant three-way interaction concerning Booker, gender, and political conservatism. This finding suggests that, on average, there was no significant reduction in sentence severity for males and females following the Booker decision, but, there were some inter-district variabilities concerning the degree to which the gender differential in sentencing outcomes actually differed across court communities. This is an interesting finding, especially because it was only revealed via the political conservatism measure. Once again, this partially confirms the argument that the two conservatism measures are differentially tied to gender equality in sentencing.

To conclude, an important take-away message from this study is the interconnectedness of judicial discretion, court communities, and gender differentials in sentencing. Because the main implication of the Booker and Gall decisions relates to the level of judicial discretion allowed in federal courts, the finding that the impact of these policy changes differed by the ecological contexts of court community adds to a growing recognition that judicial discretion does not exist in isolation. It shows that the use of discretion is embedded within local legal culture and socio-cultural environments surrounding the court communities. Ulmer's conceptualization of sentencing is quite relevant in this regard. According to Ulmer (2012, p. 8) sentencing should be conceptualized as the "joint acts produced by the discretion and interactions of judges, prosecutors, defense attorneys, and sometimes probation officers . . . [that] are embedded in [. . .] local court communities, which are in turn embedded in local socio-cultural

contexts.” Taken as a whole, the overarching conclusion of STUDY-2 is that judicial discretion cannot be detached from its social contexts.

Strengths and Limitations of the Present Study

This dissertation builds on prior research by merging two distinct bodies of literature and extends the previous studies in several significant ways. First, of paramount importance is the attempt to situate the issue of gender disparity within broader socio-political contexts. With few exceptions (see Bontrager, 2013; Helm & Jacobs, 2002), prior research on gender disparity in sentencing outcomes has been heavily focused on examining gender and gender-related factors at the individual defendant level. This dissertation, which is theoretically and methodologically sophisticated, is one of the first studies to move beyond this focus on the individual and to demonstrate that judicial decision-making in relation to gender can be dissected through a contextual lens.

Second, assessing the impact of policy changes against the backdrop of filtering is another unique contribution of this dissertation. It should be noted that there are not many studies attempting to assess the impact of sentencing reforms (Engen, 2009) and even fewer studies that assess their impact through a contextual lens (Kim et al., forthcoming), despite the voluminous body of research in the wake of the sentencing reforms. As noted, extant research on social ecology and sentencing has placed too much emphasis on showing whether and how some aggregate level theoretical constructs account for inter-district level variation in sentencing outcome (see Johnson, 2005). This dissertation sought to bring the issue of filtering to the foreground of sentencing research, thus

drawing researchers' attention to this neglected, but critical, aspect of the court community perspective.

Third, this dissertation also attempted to address recent methodological debates about appropriate strategies for modeling sentencing decisions. As acknowledged by one of the leading experts in sentencing research, Hofer (2013), there appears to be a methodological divide between two camps: criminologists (see Ulmer et al., 2011b) and empirical legal scholars (see Fischman & Schanzenbach, 2011, 2012; Starr & Rehavi, 2013). Throughout this dissertation, I sought to lay out important methodological issues, including departure and mandatory minimum controls, that the two camps do not necessarily agree upon and attempted to show that employing different model specifications produces different results. This approach is also another strength of this study, especially given the debates on methodological choices. Future studies in sentencing should incorporate this replication strategy.

Although this dissertation has significantly contributed to an understanding of the dynamics regarding sentencing outcomes, gender, ecological contexts and policy changes, it is not without limitations. A first limitation is related to the lack of pre-conviction data (Blumstein et al., 1983; Bushway & Piehl, 2007; Engen, 2009). This approach, which is a dominant commonality in sentencing research, utilized only conviction data and thus is vulnerable to a potential sample selection bias. Clearly this is a limitation that needs to be noted, as this potential censoring issue could have influenced the main conclusions of this study. From a slightly different angle, however, this limitation also offers an interesting

question to examine. Given the widespread chivalrous attitudes across the criminal justice system, it would be interesting to investigate the possibility that the social contextual variables may have distinct impacts on the different stages of criminal case processing, as courtroom actors are subject to unique constraints. For instance, prosecutors may be more susceptible to the social pressure from their environment, because they are appointed officials who are known to be more politically ambitious, as compared to the federal judges (who have lifetime appointments) in the same courtroom workgroups (Worrall, 2008).

Another important limitation of the current study concerns measurement issues; some important variables were not directly measured or were omitted. To begin, following previous work reporting a positive relationship between religious conservatism and conservative gender norms (see Moore & Vanneman, 2003), this study's measure of religious conservatism was employed as a proxy for gender norms or gender stereotypes and the overall results of this study were interpreted based on that assumption, as a variable measuring gender norms is not available at the federal district court level. Even though this measurement strategy is quite consistent with that of most prior sentencing research, the overall findings of this dissertation need to be interpreted with caution.²¹

²¹ Baumer and Martin (2013) pointed out that extant empirical literature on ecological contexts and sentencing has paid almost exclusive attention on what they called structural conditions of community and indirect measures of key variables (i.e. % Black as a proxy for perceived racial threats), thus yielding an inconclusive evidence. However, it is worthwhile to note that their study found stronger impacts of contextual variables when used as variables that directly measured the concepts. Therefore, it is also plausible that

A similar criticism could also be raised with regard to the court community perspective. Even though key theoretical arguments underlying the approach are premised on the court-organizational level, relatively little empirical attention has been devoted to directly measuring and quantifying the impact of the key concepts, such as local legal culture or going rates (Dixon, 1995). Some studies attempt to operationalize the local legal culture by employing as a proxy variable the rate of downward departure (Kautt, 2002) and the proportion of trial cases (Johnson, 2005), even though critics argued that the measures were circulatory. Like prior research, this study assumed that at least some of the impact of religious and political conservatism worked through local legal culture or substantive rationalities built into court communities. Because these variables were assumed but not measured, there is a possibility of omitted variable bias. Thus, future studies examining court environments' influence on judicial decision-making also need to take into account these measurement issues to take the research one step further.

Last, even though this dissertation is one of the first studies attempting to uncover the relationship between ecological contexts and gender disparity in sentencing using longitudinal data, future studies should conduct a more stringent test of the impact of the ecological variables on sentencing outcomes. In that regard, prior research on court communities has been predominantly based on cross-sectional studies, only showing that the ecological contexts of court community are *associated* with some types of court

the main conclusions of this dissertation would have been more substantiated, when the measure directly tapping into the gender norms had been used.

outcomes or sentencing disparity. One simple avenue to address this issue is to employ variables measuring changes in religious and political conservatism to determine whether changes in the contextual measures could account for changes in gender disparity in sentencing outcomes (see Wang & Mears, 2010b).

Theoretical and Policy Implications

Theoretical Implications

One of the intended contributions of this dissertation was to further our understanding of the nature of gender disparity in sentencing by looking at gender disparity through an ecological context. By doing so, it was expected that this dissertation would provide some important insights on why females are at an advantage over males at sentencing. Taken as a whole, the main findings of this dissertation seemed to provide support for the focal concerns perspective. This study revealed that the female discount was substantially smaller in courts embedded in communities with higher levels of religious and political conservatism, which suggests that women offenders in these court communities were more likely than women adjudicated in communities with lower levels of political and religious conservatism to be perceived as blameworthy. Because the opposite scenario did fail to receive empirical support, I conclude that the focal concerns perspective prevails over the social control argument proposed by Kruttschnitt (1982, 1984) in accounting for gender disparity in sentencing outcomes.

This dissertation also provides support for the argument that blameworthiness and practical constraints/social costs are the two most important elements of the focal concerns perspective (Steffensmeier et al., 1993). This interpretation is most evident with the finding related to the impact of religious conservatism on gender disparity. As noted earlier, whereas the sentences imposed on male offenders, especially (young) African American male offenders, were not influenced by the level of religious conservatism, the sentences given to female offenders were shaped by the level of religious conservatism. This suggests that religious conservatism triggers the domain of blameworthiness through conservative gender norms, thus enhancing attributions of blameworthiness for female offenders in religiously conservative districts. At the same time, it may be the case that the impact of religious conservatism did not affect the domain of dangerousness, argued to be strongly associated with singling out African American male offenders for harsher punishments (Steffensmeier et al., 1998), such that the male offenders in those districts with higher levels of religious conservatism were not viewed as particularly dangerous to the communities.

The results of this dissertation also provide an important insight on the interconnectedness of the focal concerns and court community perspectives. Recent scholarship has argued that focal concerns could be situated within the court community perspective, as the interpretation and prioritization of the focal concerns is shaped by the characteristics of court communities (Kramer & Ulmer, 2008; Ulmer, 2012). In that regard, this dissertation offers full support for two main hypotheses deriving from the court community perspective. This study found, first, that both average sentence severity

and the treatment of female offenders significantly differed across federal district courts; this adds to a growing body of evidence that social contexts matter. Second, this dissertation provides support for the hypothesis that the impact of policy changes is always filtered through the substantive rationality of the court community (Einstein et al., 1988). It was found that the impact of the federal guidelines' transition into an advisory status had distinct impacts on both sentencing outcomes in general and gender disparity in particular.

Even though it is true that this dissertation documented evidence in support of the utilities of two main theoretical perspectives in judicial decision-making, some comments about the limitations of these two perspectives are also warranted. One glaring limitation concerning the focal concerns theory is the fact that there is no single empirical study which provides direct evidence to support the propositions. All the available evidence, including those from this dissertation, is at most indirect. Future studies should address this key limitation by conducting studies designed to directly measure judges' perceptions of each element of focal concerns and the way in which these concerns influence their sentencing decisions.

The court community perspective also has limitations, the first of which concerns the lack of specificity of key theoretical constructs subsumed under the perspective. For instance, the local legal culture and the going rate constitute two of the most important concept in the court community perspective. But the theoretical framework fails to offer any guidance on how to conceptualize and quantify these concepts and, for this reason,

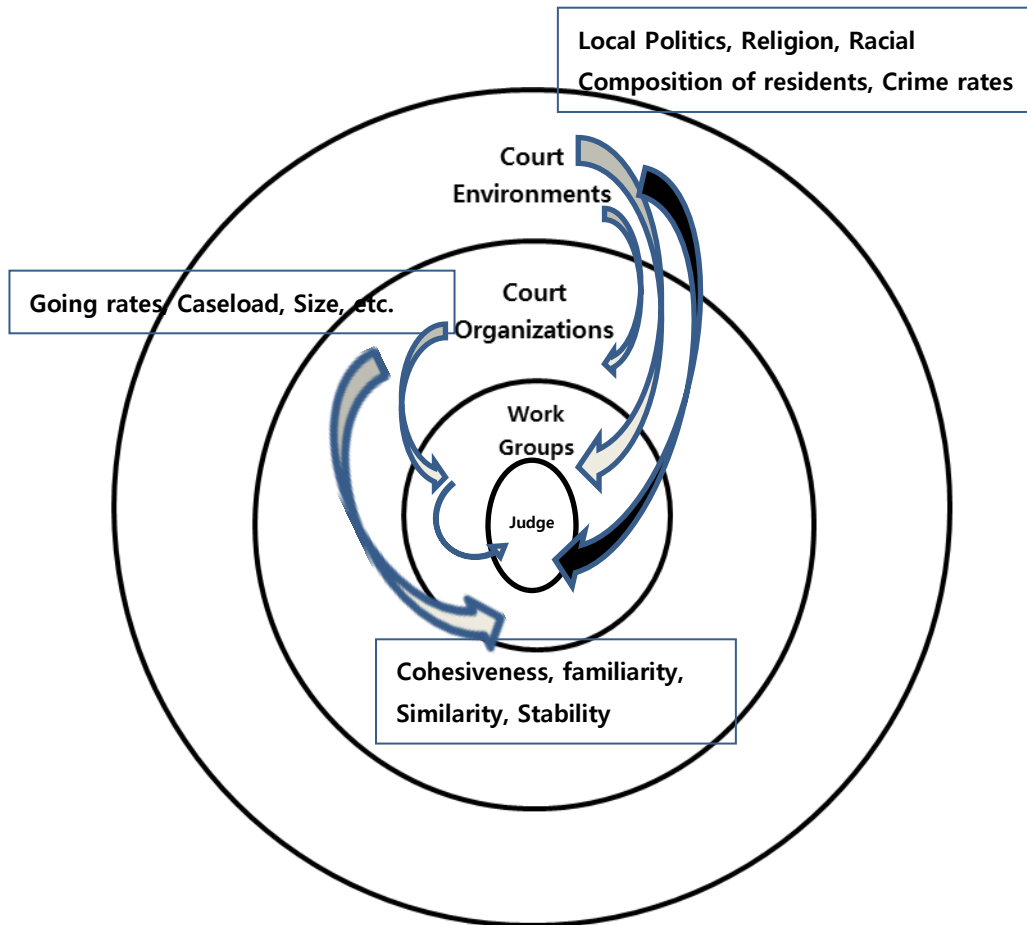
many empirical studies following in this tradition fail to incorporate these concepts in their empirical inquiries. What we know at this moment is that sentencing outcomes and processes vary across court communities and some theoretical constructs account for the variation. Therefore, a fair assessment of the current body of evidence concerning the main arguments of court community perspective may be not as strong as the theory argues.

Another key limitation concerning the court community perspective lies in the fact that the perspective puts more emphasis on the court-organizational level explanations, with the influence of court environmental factors receiving secondary attention (Myers & Talarico, 1987). But the empirical research following this perspective has devoted dominant attention to the environmental factors with the organizational explanations receiving little or no empirical interest. Thus there appears to be a noticeable gap between the theory and the empirical research. Against this backdrop, this dissertation suggests the need to expand and elaborate on the core concepts of the court community perspective. This dissertation found that at least some impact of religious and political conservatism worked through the local legal culture or substantive rationality inculcated in court communities. However, these key measures were not included in the empirical models, but were just assumed.

Another important limitation also concerns the fact that there is a lack of theoretical sophistication on the connections between the organizational and the environmental levels within which courts operate. Most importantly, our understanding

of the potential mechanisms by which court environmental factors influence judicial decision-making is under-developed. As can be seen in Figure-13, sentencing is a complex process that is influenced by multiple sources from many different levels. Yet, our understanding and empirical inquiries on court environmental factors have focused on the direct impact of those factors on judicial decision-making. As discussed previously, some of the important effects of environmental variables, such as religious and political conservatism, work through other important layers surrounding individual judges and other courtroom actors. To put it differently, substantial portions of the environmental effect may be indirect, much of which is already controlled for. For instance, the impact of religious and political conservatism observed in this dissertation is the one represented by the black shaded pathway in Figure-13 and many other indirect pathways were ignored in the analysis. Theorizing and examining the pathways questions would also require conceptualizing this question through a temporal dimension (Myers & Talarico, 1987). Whereas some of the environmental influences may shape judicial decision-making immediately, for instance, political conservatism in this dissertation, other social factors, such as religious conservatism in this study, may take a while to exert an impact on judicial decision-making. Extant literature stay silent on this issue, only suggesting that court environmental factors matter.

Figure-13 Multiple layers surrounding federal judges



Policy Implications

The findings from this dissertation have several policy implications. One of the fundamental policy implications concerns the issue of gender neutrality in sentencing outcomes. As noted previously, there is no clear consensus as to whether gender differentials require some type of reform. This issue starts with the way in which gender equality in sentencing is conceptualized. What many commentators find problematic is

the concept of a *restricted notion of equality* (Daly & Tonry, 1997). In order to better appreciate this concept, it is helpful to pay attention to a quote from Chesney-Lind and Pollock (1995, p. 170):

In order for women offenders to receive justice, it must be recognized that men and women inhabit different social realities and that women are not necessarily best served if they are treated in ways that assume their needs are identical to their male counterparts.

Although debatable, I argue that there should be a serious reconsideration of gender equality in sentencing. As Raeder (1993) suggests, if the main goal of sentencing is justice, the special circumstances surrounding women should be taken into account, rather than completely ignored. In many cases, the issue concerning gender neutrality eventually boils down to what is called a gender-related factor, such as pregnancy, being a single mother, and other family ties issues. In order to achieve the goal of reducing *unwarranted* gender disparity and give justice to women offenders, I argue that rendering gender-related factors legitimate considerations should be an important first step. What many reformers found problematic in discussing gender disparity during the sentencing reform movement was the female discount mainly resulting from gender stereotypes (Williams, 1982) and there are many commentators who claim that departures from the guidelines based on family ties are fully justified (Raeder, 1993, 2005; Daly, 1987).

Another important but more detailed point to be raised when discussing gender equality in sentencing involves how to achieve it from a policy perspective. I argue that overall sentence severity should play a critical role in devising sentencing guidelines.

Roughly speaking, there are two potential ways by which gender equality in sentencing could be achieved in terms of severity of punishment. A first solution would be to increase the level of punishment imposed on female offenders to that imposed on male offenders; the other possibility would be to decrease the level of punishment meted out to male offenders. The federal sentencing guidelines took the former approach to an extreme level, which resulted in quite counter-intuitive findings—an *increase* in gender disparity in the wake of the implementation of the guidelines (USSC, 2004). Thus, if the goal of any reform movement is to reduce gender disparity in sentencing outcomes, policymakers should learn this important lesson by placing overall sentence severity at the center of the discussion.

Turning to the issue of inter-jurisdiction variation in sentencing outcomes, just as there is no clear consensus on whether gender disparity is warranted, there are two differing views on the appropriateness of inter-district disparities in sentencing outcomes. As Ulmer and Johnson (2004) put it, these disparities present a dilemma concerning two fundamental issues underlying the American society and criminal justice system. On the one hand, the mere fact that the level of punishment depends on the place where a defendant is sentenced undermines the principle of equal justice. By contrast, the emphasis put on decentralized decision-making and local autonomy also permits the use of localized or substantive justice developed within specific local jurisdictions (see also Johnson, Ulmer, and Kramer, 2008). This issue has taken on added importance in federal district courts because of their nationwide jurisdiction. So, even though it is true that inter-jurisdictional variation in sentencing outcomes was one of the motivating factors for

the sentencing reform movement in federal courts (Breyer, 1988-1989), Congress and policymakers in federal courts should recognize that federal courts are arenas in which extremely complex layers of environmental factors come into play and the reform movement designed to increase the level of consistency concerning inter-jurisdictional variation is challenging

Finally, this dissertation has demonstrated that female offenders are treated differently from jurisdiction to jurisdiction and that the impact of the transition of federal guidelines into an advisory status also created district-specific results. These results may have been attributable to the lower levels of consensus by court actors across federal district courts as to what extent the gender and inter-jurisdictional variations are problematic. In addition, it is also important to note that the development and implementation of the federal sentencing guidelines was too radical in terms of constraining judicial discretion and elevating the level of punishment. Partly, this may have contributed to the varying results. And it is important to remember that, as Einstein et al. (1988, p. 294) argued, “the more radical a proposed change, the less likely is its adoption.”

Conclusion

The purpose of this dissertation was to view gender disparity in sentencing outcomes through ecological lenses. The findings demonstrate that the gender differential in sentencing outcomes is indeed contextualized by religious and political conservatism. The main findings of this dissertation add to a growing body of research recognizing the

complex influences of court environments and call for more follow-up studies specifically focusing on the relationship between gender and sentencing through ecological contexts. Taken together, the findings of this dissertation reveal that judicial decision-making is influenced by multiple layers of court environments, which suggests that the narrow focus on immediate circumstances concerning gender and gender-related factors should be accompanied by a newer approach emphasizing the ecological contexts of court communities.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage.
- Albonetti, C. A. (1991). An integration of theories to explain judicial discretion. *Social Problems*, 38, 247–266.
- Albonetti, C. A. (1997). Sentencing under the federal sentencing guidelines: effects of offender characteristics, guilty pleas, and departures on sentence outcomes for drug offenses, 1991–1992. *Law and Society Review*, 31, 789–822.
- Albonetti, C. A. (1998). The Role of Gender and Departures in the Sentencing of Defendants Convicted of White-collar Offense Under the Federal Sentencing Guidelines. In Jeffery Ulmer, ed., *Sociology of Crime, Law, and Deviance*. Vol. 1. Greenwich: JAI Press.
- Albonetti, C. A. (2002). Joint Conditioning Effect of Defendant's Gender and Ethnicity on Length of Imprisonment under the Federal Sentencing Guidelines for Drug Trafficking/Manufacturing Offenders, The. *Journal of Gender, Race, and Justice*, 6, 39-60.
- Allen, F. A. (1964). *The borderland of criminal justice: Essays in law and criminology* (pp. 25-41). Chicago: University of Chicago Press.
- Allen, H. (1987). *Justice unbalanced: Gender, psychiatry, and judicial decisions* (p. 28). Milton Keynes: Open University Press.
- Alexander, M. (2012). *The new Jim Crow: Mass incarceration in the age of colorblindness*. The New Press.
- Allison, P. D. (2000). Multiple imputation for missing data: A cautionary tale. *Sociological Methods and Research*, 28, 301-309.
- American Statistical Association of Religious Bodies. (2002). *Religious congregations and membership in the United States 2000* [CD-ROM]. Nashville, TN: Glenmary Research Center.
- Anderson, E. A. (1976). The chivalrous treatment of the female offender in the arms of the criminal justice system: a review of the literature. *Social Problems*, 23, 350-357.
- Anderson, A. L., & Spohn, C. (2010). Lawlessness in the federal sentencing process: A test for uniformity and consistency in sentence outcomes. *Justice Quarterly*, 27, 362-393.

- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of personality and social psychology*, 71, 230-258.
- Bartkowski, J. P. (1997). Debating patriarchy: Discursive disputes over spousal authority among evangelical family commentators. *Journal for the Scientific Study of Religion*, 31, 15-31.
- Bartkowski, J. P., & Read, J. G. (2003). Veiled submission: Gender, power, and identity among evangelical and Muslim women in the United States. *Qualitative Sociology*, 26, 71-92.
- Baumer, E. P. (2013). Reassessing and redirecting research on race and sentencing. *Justice Quarterly*, 30, 231-261.
- Baumer, E. P., & Martin, K. H. (2013). Social Organization, Collective Sentiment, and Legal Sanctions in Murder Cases. *American Journal of Sociology*, 119, 131-182.
- Beckett, K. (1997). *Making Crime Pay: Law and Order in Contemporary American Politics*. New York: Oxford University Press.
- Beckett, K., & Sasson, T. (2003). *The politics of injustice: Crime and punishment in America*. Sage Publications.
- Beckett, K. & Western, B. (2001). Governing social marginality: Welfare, incarceration, and the transformation of state policy. *Punishment and Society*, 3, 43-59.
- Belknap, J. (2001). *The Invisible Woman: Gender, Crime and Justice*, 2nd ed. Boulder: University of Colorado Press.
- Berk, R. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, 48, 386-98.
- Berman, D. A. (2005). Assessing federal sentencing after Booker. *Federal Sentencing Reporter*, 17, 291-294.
- Bickle, G. S., & Peterson, R. D. (1991). The impact of gender-based family roles on criminal sentencing. *Social Problems*, 38, 372-394.
- Black, D. (1976). *The Behavior of Law*. NY: Academic.
- Blackwell, B. S., Holleran, D., & Finn, M. A. (2008). The impact of the Pennsylvania sentencing guidelines on sex differences in sentencing. *Journal of Contemporary Criminal Justice*, 24, 399-418.
- Blalock, H. M. (1967). *Toward a Theory of Minority Group Relations*. New York: Wiley.

- Blanchard, D. A., & Prewitt, T. J. (1993). *Religious violence and abortion: The Gideon Project*. Gainesville: University of Florida Press.
- Blumer, H. (1955). Attitudes and the social act. *Social Problems*, 3, 59-65.
- Blumstein, A., Cohen, J., Martin, S.E., & Tonry, M.H. (1983). Research on sentencing: The search for reform. Vol. 1. Washington, DC: National Academy Press.
- Boles, Jacqueline & Maxine P. Atkinson. (1986). Ladies: South by Northwest. *Sociological Spectrum*, 6, 63-81.
- Bonilla-Silva, E. (1997). Rethinking racism: Toward a structural interpretation. *American Sociological Review*, 465-480.
- Bonilla-Silva, E. (2000). This is a white country. The racial ideology of the western nations of the world system. *Social Inquiry*, 70, 188-214.
- Bontrager, S., Bales, W., & Chiricos. T. (2005). Race, ethnicity, threat and the labeling of convicted felons. *Criminology*, 43, 589-622.
- Bontrager, S., Barrick, K., & Stupi, E. (2013). Gender and Sentencing: A Meta-Analysis of Contemporary Research. *Journal of Gender, Race and Justice*, 16, 349–372.
- Boritch, H. (1992). Gender and Criminal Court Outcomes: A Historical Analysis. *Criminology*, 30, 293-326.
- Bowman, F. O. III. (2006). The Year of Jubilee... or Maybe Not: Some Preliminary Observations About the Operation of the Federal Sentencing System After Booker. *Houston Law Review*, 43, 281.
- Bowman, F. O., III. (2012). Nothing is not enough: Fix the absurd post-booker federal sentencing system. *Federal Sentencing Reporter*, 24, 356-368.
- Braithwaite, J., & Pettit, P. (1990). Not just deserts: A republican theory of criminal justice (Vol. 78). Oxford: Clarendon Press.
- Brannon, L. (1996). *Gender: psychological perspectives*. Boston, MA: Allyn and Bacon.
- Brennan, P. K. (2006). Sentencing female misdemeanants: An examination of the direct and indirect effects of race/ethnicity. *Justice Quarterly*, 23, 60–95.
- 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, 200-217.
- Breyer, S. (1988). The Federal Sentencing Guidelines and the Key Compromises Upon Which They Rest. *Hofstra Law Review*, 17, 1-50.

- Brinkerhoff, M. B., Jeffrey, C. J., & Mackie, M. M. (1987). Mormonism and the Moral Majority Make Strange Bedfellows? An Exploratory Critique. *Review of Religious Research*, 28, 236–51.
- Britt, C. L. (2000). Social context and racial disparities in punishment decisions. *Justice Quarterly*, 17, 707-732.
- Bushway, S., & Forst, B. (2013). Studying discretion in the processes that generate criminal justice sanctions. *Justice Quarterly*, 30, 199-222.
- Bushway, S., Johnson, B.D., & Slocum, L.A. (2007). Is the magic still there? The use of the Heckman two-step correlation for selection bias in criminology. *Journal of Quantitative Criminology*, 23, 151-178.
- Bushway, S., & Piehl, A. (2001). Judging judicial discretion: Legal factors and racial discrimination in sentencing. *Law & Society Review*, 35, 733-764.
- Cameron, C., & Trivedi, P. K. (2010). *Microeconometrics Using Stata, Revised Edition*. College Station: Tex.: Stata Press.
- Caravelis, C., Chiricos, T., & Bales, W. (2011). Static and dynamic indicators of minority threat in sentencing outcomes: A multi-level analysis. *Journal of Quantitative Criminology*, 27, 405-425.
- Carson, E.A., & Golinelli, D. (2013). *Prisoners in 2012*. Washington, D.C. Bureau of Justice Statistics.
- Carter, S. & Borch, C. (2005). Assessing the effects of urbanism and regionalism on gender role attitudes, 1974-1998. *Sociological Inquiry*, 75, 548-563.
- Castro, D. O. (1998). Hot blood and easy virtue': Mass media and the making of racist latino/a stereotypes. *Images of color, images of crime*.
- Chambliss, W. J. (1964). A sociological analysis of the law of vagrancy. *Social Problems*, 67-77.
- Chambliss, W. J. (1994). Policing the ghetto underclass: The politics of law and law enforcement. *Social Problems*, 41, 177-198.
- Chambliss, W. J., & Seidman, R. B. (1971). *Law, order, and power*. Reading, MA: Addison-Wesley.
- Chesney-Lind, M. (1991). Patriarchy, prisons, and jails: A critical look at trends in women's incarceration. *The Prison Journal*, 71, 51-67.
- Chesney-Lind, M. (1995). Rethinking Women's Imprisonment. Pp .105-17 in *The Criminal Justice System and Women: Offenders, Victims, and Workers*. 2nd ed., edited by B. R. Price and N. J. Sokoloff. New York: McGraw-Hill.

- Chesney-Lind, M. (1997). *The Female Offender: Girls, Women, and Crime*. Thousand Oaks: SAGE Publications.
- Chesney-Lind, M. (1998). Women in Prison: From Partial Justice to Vengeful Equity. *Corrections Today*, 60, 66-71.
- Chesney-Lind, M. (2002). Criminalizing victimization: The unintended consequences of pro-arrest policies for girls and women. *Criminology & Public Policy*, 2, 81-90.
- Chesney-Lind, M. & Pollock, J. M. (1995). Women's Prisons: Equality with a Vengeance. Pp. 155-175 in *Women, Law, and Social Control*, ed. Alida V. Merlo and Joycelyn M. Pollock. Allyn and Bacon: Boston.
- Cherlin, A. J. (1992). Going to extremes: Family structure, children's well-being, and social science. *Demography*, 36, 421-428.
- Chiricos, T. G., & Bales, W. D. (1991). Unemployment and punishment: An empirical assessment. *Criminology*, 29, 701-724.
- Chiricos, T. G., & Delone, M. A. (1992). Labor surplus and punishment: A review and assessment of theory and evidence. *Social Problems*, 39, 421-446.
- Cott, N. F. (1997). *The Bonds of Womanhood: "Woman's Sphere" in New England, 1780-1835*. Yale University Press.
- Crawford, C. (2000). Gender, race, and habitual offender sentencing in Florida. *Criminology*, 38, 263-280.
- Crawford, C. Chiricos, T. & Kleck, G. (1998). Race, racial threat, and sentencing of habitual offenders. *Criminology*, 36, 481-511.
- Crew, K. B. (1991). Sex differences in criminal sentencing: Chivalry or patriarchy? *Justice Quarterly*, 8, 59-84.
- Cullen, F. T. & Jonson, C. L. (2011). *Correctional theory: Context and consequences*. Sage.
- Curran, D. (1983). Judicial discretion and defendant's sex. *Criminology*, 21, 41-58.
- Curry, T. R., Lee, G., & Rodriguez, S. F. (2004). Does victim gender increase sentence severity? Further explorations of gender dynamics and sentencing outcomes. *Crime & Delinquency*, 50, 319-343.
- Daly, K. (1987a). Structure and Practice of familial-based justice in a criminal court. *Law and Society Review*, 21, 267-290.
- Daly, K. (1987b). Discrimination in the criminal courts: family, gender, and the problem of equal treatment. *Social Forces*, 66, 152-175.

- Daly, K. (1989a). Rethinking judicial paternalism: Gender, workfamily relations, and sentencing, *Gender & Society*, 3, 9–36.
- Daly, K. (1989b). Neither Conflict nor Labeling nor Paternalism Will Suffice: Intersections of Race, Ethnicity, Gender, and Family in Criminal Court Decisions. *Crime and Delinquency*, 35, 136-168.
- Daly, K. (1994). *Gender, crime, and punishment*. New Haven, CT: Yale University Press.
- Daly, K., & Bordt, R. L. (1995). Sex effects and sentencing: An analysis of the statistical literature. *Justice Quarterly*, 12, 141–175.
- Daly, K., & Tonry, M. (1997). Gender, race, and sentencing. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 22, pp. 201–252). Chicago, Illinois: University of Chicago Press.
- Davis, K. C. (1969). *Discretionary justice: A preliminary inquiry*. LSU Press.
- Demuth, S. (2003). Racial and ethnic differences in pretrial release decisions and outcomes: A comparison of Hispanic, black, and white felony arrestees. *Criminology*, 41, 873-908.
- DiPrete, T. A., & Grusky, D. B. (1990). The multilevel analysis of trends with repeated cross-sectional data. *Sociological Methods*, 20, 337-368.
- Dixon, J. (1995). The organizational context of criminal sentencing. *American Journal of Sociology*, 100, 1157-1198.
- Doerner, J., & Demuth, S. (2010). The independent and joint effects of race/ethnicity, gender, and age on sentencing outcomes in U.S. Federal courts. *Justice Quarterly*, 27, 1–27.
- Dunworth, T., & Weisselberg, C. D. (1992). Felony cases and the federal courts: The Guidelines experience. *Southern California Law Review*, 66, 99-153.
- Durkheim, É. (1966). *The Division of Labor in Society*. New York: Free Press.
- Ehrlich, E. (2002). *Fundamental Principles of the Sociology of Law*. New Brunswick, NJ: Transaction Publishers.
- Eisenstein, J., Flemming, R. B., & Nardulli, P. (1988). *The contours of justice: Communities and their courts*. Boston, MA: Little, Brown.
- Eisenstein, J., & Jacob, H. (1977). *Felony justice: An organizational analysis of criminal courts*. Boston, MA: Little, Brown
- Engen, R. L. (2009). Assessing determinate and presumptive sentencing: Making research relevant. *Criminology & Public Policy*, 8, 323-335.

- Engen, R. L. (2011). Racial disparity in the wake of Booker/Fanfan making sense of “messy” results and other challenges for sentencing research. *Criminology & Public Policy*, *10*, 1139-1149
- Engen, R.L., & Gainey, R.R. (2000a). Conceptualizing legally relevant factors under guidelines: a reply to Ulmer. *Criminology*, *38*, 1245–1252.
- Engen, R.L., & Gainey, R.R. (2000b). Modeling the effects of legally relevant and extralegal factors under sentencing guidelines: the rules have changed. *Criminology*, *38*, 1207–1230.
- Engen, R. L., & Steen, S. (2000). The power to punish: Discretion and sentencing reform in the war on drugs. *American Journal of Sociology*, *105*, 1357-1395.
- Farnworth, M & Horan, P. (1980). Separate Justice: An analysis of race differences in court processes. *Social Science Research*, *9*, 381-399.
- Farnsworth, M., & Teske, R. H. C., Jr. (1995). Gender differences in filling court processing: Testing three hypotheses of disparity. *Women and Criminal Justice*, *6*, 23-44.
- Farrell, A., & Ward, G. (2011). Examining district variation in sentencing in the post-Booker period. *Federal Sentencing Reporter*, *23*, 318-325.
- Fearn, N. E. (2005). A multilevel analysis of community effects on criminal sentencing. *Justice Quarterly*, *22*, 452-22, 487.
- Feldmeyer, B., & Ulmer, J. T. (2011). Racial/ethnic threat and federal sentencing. *Journal of Research in Crime and Delinquency*, *48*, 238-270.
- Fine, G. A. (1984). Negotiated orders and organizational cultures. *Annual Review of Sociology*, *10*, 239-262.
- Fischman, J. B., & Schanzenbach, M. M. (2011). Do standards of review matter? The case of federal criminal sentencing. *The Journal of Empirical Legal Studies*, *40*, 405-437.
- 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*, 729-764.
- Fishman, L. T. (1998). The black bogeyman and white self-righteousness. In C. R. Mann & M. S. Zatz (Eds.), *Images of color, images of crime*. Los Angeles: Roxbury.
- Flemming, R., Nardulli, P., & Eisenstein, J. (1992). *The craft of justice: Work and politics in criminal court communities*. Philadelphia, PA: University of Pennsylvania Press.

- Foucault, M. (1977) *Discipline and punish: The births of the prison*. London: Penguin.
- Frankel, M. E. (1972). Lawlessness in sentencing. *University of Cincinnati Law Review*, 41, 1-54.
- 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, 279-290.
- Freeborn, B. A., & Hartmann, M. E. (2010). Judicial discretion and sentencing behavior: Did the Feeney amendment rein in district judges? *Journal of Empirical Legal Studies*, 7, 355-378.
- Freiburger, T. L. (2010). The Effects of Gender, Family Status, and Race on Sentencing Decisions. *Behavioral Sciences and the Law*, 28, 378-395.
- Garland, D. (1990). *Punishment and modern society*. Oxford, UK: Clarendon Press.
- Garland, D. (2001). *The Culture of Control: Crime and Social Order in Contemporary Society*. Chicago: University of Chicago Press.
- Gilkes, C. T. (1983). From slavery to social welfare: Racism and the control of black women. In A. Swerdlow & H. Lessinger (Eds.), *Class, race, and sex: The dynamics of control*. Boston: G. K. Hall. Granovetter, 1985
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, 74, 1464-1480.
- Griffin, T., & Wooldridge, J. (2006). Sex-based disparities in felony dispositions before versus after sentencing reform in Ohio. *Criminology*, 44, 893-923.
- Gross, N., Medvetz, T., & Russell, R. (2011). The Contemporary American Conservative Movement. *Annual Review of Sociology*, 37, 325-354.
- Hagan, J. (1974). Extra-legal attributes and criminal sentencing: An assessment of a sociological viewpoint. *Law & Society Review*, 8, 357-383.
- Hagan, J., & Bumiller, K. (1983). Making Sense of Sentencing: A Review and Critique of Sentencing Research. In Al Blumenstein, Jacqueline Cohen, Susan E. Martin, and Michael Tonry (eds.), *Research on Sentencing: Search for Reform*. Vol. II. Washington, DC: National Academy of Sciences.
- Hagan, J., & Dinovitzer, R. (1999). Collateral Consequences of Imprisonment for Children, Communities, and Prisoners. *Crime and Justice: A Review of Research*, 26, 121-162.

- Hagan, J., Nagel, I. H., & Albonetti, C. (1980). The differential sentencing of white-collar offenders in ten federal district courts. *American Sociological Review*, 45, 802-820.
- Harris, R. J., & Firestone, J. M. (1998). Changes in predictors of gender role ideologies among women: A multivariate analysis. *Sex Roles*, 38, 239-252.
- Hart, H. L. A. (1968). *Punishment and responsibility: Essays in the philosophy of law*. Oxford University Press.
- Hartley, R. D., Maddan, S., & Spohn, C. (2007). Prosecutorial discretion: An examination of substantial assistance departures in federal crack-cocaine and powder-cocaine cases. *Justice Quarterly*, 24, 383-407.
- Healey, J. F. (1997). *Race, ethnicity, and gender in the United States: Inequality, group conflict, and power*. Thousand Oaks, CA: Pine Forge Press.
- Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, 5, 475-492.
- Helms, R. (2009). Modeling the politics of punishment: A conceptual and empirical analysis of 'law in action' in criminal sentencing. *Journal of Criminal Justice*, 37, 10-20.
- Helms, R., & Jacobs, D. (2002). The political context of sentencing: An analysis of community and individual determinants. *Social Forces*, 81, 577-604.
- Herzog, S., & Oreg, S. (2008). Chivalry and the moderating effect of ambivalent sexism: Individual differences in crime seriousness judgments. *Law & Society Review*, 42, 45-74.
- Hofer, P. J. (2006). Immediate and long-term effects of United States v. Booker: More discretion, more disparity, or better reasoned sentences? *Arizona State Law Journal*, 38, 425-468.
- Hofer, P. J. (2007). United States v. Booker as a natural experiment: Using empirical research to inform the federal sentencing policy debate. *Criminology & Public Policy*, 6, 433-460.
- Hofer, P. J., Blackwell, K. R., & Ruback, R. B. (1999). The effect of the federal sentencing guidelines on inter-judge disparity. *Journal of Criminal Law and Criminology*, 90, 239-321.
- Holleran, D., & Spohn, C. (2004). On the use of the total incarceration variable in sentencing research. *Criminology*, 42, 211-240.

- Horton, N. J. & Lipsitz, S. R. (2001). Multiple imputation in practice: Comparison of software packages for regression models with missing variables. *The American Statistician*, 55, 244-254.
- Huang, W.S. Wilson, Finn, R., Ruback, & Friedmann, R. R. (1996). Individual and contextual influences on sentence lengths: Examining political conservatism. *The Prison Journal*, 76, 398-419.
- Huckerby, J. (2003). Women Who Kill Their Children: Case Study and Conclusions Concerning the Differences in the Fall from Maternal Grace by Khoua Her and Andrea Yates. *Duke J. Gender L. & Pol'y*, 10, 149.
- Hurlbert, J. S. (1988). The Southern region: A test of the hypothesis of cultural distinctiveness. *Sociological Quarterly*, 30, 245-266.
- Jackson, P. I. (1989). *Minority Group Threat, Crime, and Policing*. New York: Praeger.
- Jacobs, D., & Carmichael, J. T. (2001). The politics of punishment across time and space: A pooled time-series analysis of imprisonment rates. *Social Forces*, 80, 61-89.
- Jacobs, D., & Helms, R. E. (1996). Toward a political model of incarceration: a time-series examination of multiple explanations for prison admission rates. *American Journal of Sociology*, 323-357.
- Johnson, B. D. (2005). Contextual disparities in guideline departures: Courtroom social contexts, guidelines compliance, and extralegal disparities in criminal sentencing. *Criminology*, 43, 761-796.
- Johnson, B. D. (2006). The multilevel context of criminal sentencing: Integrating judge- and county-level influences. *Criminology*, 44, 259-298.
- Johnson, B. D. (2010). Multilevel analysis in the study of crime and justice. In *Handbook of Quantitative Criminology*, eds. Alex Piquero and David Weisburd. New York: Springer Publishing.
- 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., Ulmer, J.T., & Kramer, J.H. (2008). The social context of guidelines circumvention: the case of federal district courts. *Criminology*, 46, 737-783.
- Karlsen, C. (1998). *The Devil in the Shape of a Woman*. New York: W.W. Norton and Company.

- Kaukinen, C. E. (1995). Women lawbreakers constructed in terms of traditional definitions of femininity: The sentencing of women in conflict with the law. Unpublished master's thesis. University of Windsor, Ontario, Canada.
- Kautt, P. M. (2002). Location, location, location: Interdistrict and intercircuit variation in sentencing outcomes for federal drug-trafficking offenses. *Justice Quarterly*, 19, 633-671.
- Kautt, P. M., & Spohn, C. (2002). Crack-ing down on black drug offenders? Testing for interactions between offender race, drug type, and sentencing strategy in federal drug sentences. *Justice Quarterly*, 19, 1-36.
- Kim, B., Cano, M. V., Kim, K., & Spohn, C. (forthcoming). The Impact of United States v. Booker and Gall/Kimbrough v. United States on Sentence Severity Assessing Social Context and Judicial Discretion. *Crime & Delinquency*, doi: 10.1177/0011128714551783
- King, R. D. (2008). Conservatism, Institutionalism, and the Social Control of Intergroup Conflict. *American Journal of Sociology*, 113, 1351-1393.
- King, G., Honaker, J., Joseph, A., & Scheve, K. (2001). Analyzing incomplete political science data: An alternative algorithm for multiple imputation, *American Political Science Review*, 95, 49-69.
- Knapp, K. A. (1984) The impact of the Minnesota sentencing guidelines: Three year evaluation. Minnesota Sentencing Guidelines Commission.
- Klein, M. W., (1979), Deinstitutionalization and Diversion of Juvenile Offenders: A Litany of Impediments. Pp. 145-201 in Crime and Justice: An Annual Review of Research, vol. 1, ed. By Norval Morris and Michael Tonry. The University of Chicago Press: Chicago,IL.
- Klein, D. & J. Kress. (1976). Any woman's blues: a critical overview of women, crime and the criminal justice system. *Crime and social justice*. 5, 34-49.
- Koons-Witt, B. (2002). The effect of gender on the decision to incarcerate before and after the introduction of sentencing guidelines. *Criminology*, 40, 297-328.
- Kramer, J. H., & Ulmer, J. T. (1996). Sentencing Disparity and Departures From Guidelines. *Justice Quarterly*, 13, 81-105.
- Kramer, J. H., & Ulmer, J. T. (2002). Downward departures for serious violent offenders: Local court 'corrections' to Pennsylvania's sentencing guidelines. *Criminology*, 40, 897-932.
- Kramer, J. H. & Ulmer, J. T. (2009). *Sentencing Guidelines: Lessons from Philadelphia*. Boulder, CO: Lynne Rienner Publishers.

- Kruttschnitt, C. (1981). Social Status and Sentences of Female Offenders. *Law & Society Review*, 15, 247-265.
- Kruttschnitt, C. (1982). Women, crime, and dependency: An application of the theory of law. *Criminology*, 19, 495–513.
- Kruttschnitt, C. (1984). Sex and Criminal Court Dispositions: The Unresolved Controversy. *Journal of Research in Crime and Delinquency*, 21, 213-232.
- Kruttschnitt, C. & Green, D. (1984). The Sex-Sanctioning Issue: Is it History? *American Sociological Review*, 49, 541-551.
- Kruttschnitt, C. & Gartner, R. (2003). Women's imprisonment. *Crime and Justice*, 1-81.
- Kruttschnitt, C. & McCarthy, D. (1985). Familial social control and pretrial sanctions: Does sex really matter? *The Journal of Criminal Law and Criminology*, 76, 151-175.
- Kruttschnitt, C., & Savolainen, J. (2009). Ages of Chivalry, Places of Paternalism Gender and Criminal Sentencing in Finland. *European Journal of Criminology*, 6, 225-247.
- Levin, M. A. (1972). Urban politics and judicial behavior. *The Journal of Legal Studies*, 193-221.
- Levitt, S. D. (2004). Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not. *Journal of Economic perspectives*, 163-190.
- Lewin, M. (1984). In the shadow of the past: Psychology portrays the sexes. New York: Columbia University Press.
- Little, R. J. A. (1992). Regression with missing X's: A review. *Journal of the American Statistical Association*, 87, 1227–1237.
- Lipton, D., Martinson, R., & Wilks, J. (1975). The effectiveness of correctional treatment: A survey of treatment evaluation studies. New York: Praeger.
- Lubitz, R. L., & Ross, T. W. (2001). *Sentencing guidelines: Reflections on the future*. Washington, DC: US Department of Justice, Office of Justice Programs, National Institute of Justice.
- Lukies, G. (1971). History and class consciousness. Trans Rodney Livingstone. Cambridge: MIT Press.
- Lupfer, M., Hopkinson, P.L. & Kelly, P. (1988). An Exploration of the Attributional Styles of Christian Fundamentalists and of Authoritarians. *Journal for the Scientific Study of Religion*, 27, 389–398.

- Madriz, E. I. (1997). Images of Criminals and Victims a Study on Women's Fear and Social Control. *Gender & Society, 11*, 342-356.
- Martinson, R. (1974). What works? Questions and answers about prison reform. *The public interest, 35*, 22-54.
- Martin, C. L., & Little, J. K. (1990). The Relation of Gender Understanding to Children's Sex-typed Preferences and Gender Stereotypes. *Child Development, 61*, 1427-1439.
- Mauer, M. (2011). Sentencing reform: Amid mass incarcerations-guarded optimism. *Crime and Justice, 26*, 27.
- Mauer, M. (2013). The changing racial dynamics of women's incarceration. Washington, DC: Sentencing Project.
- Mauer, M., & Chesney-Lind, M. (2002). *Invisible Punishment: The Collateral Consequences of Mass Imprisonment*. New York, NY: New Press.
- Mauer, M., & King, R.S. (2007). Uneven justice: State rates of incarceration by race and ethnicity. Washington, DC: The Sentencing Project.
- Mauer, M., Potler, C., & Wolf, R. (1999). Gender and justice: Women, drugs, and sentencing policy. Washington, DC: Sentencing Project.
- Maxfield, L. D., & Kramer, J. H. (1998). Substantial Assistance: An Empirical Yardstick Gauging Equity in Current Federal Policy and Practice. *Federal Sentencing Reporter, 11*, 6.
- McIntosh, W. V. (1991). Courts and socioeconomic change. *The American Courts: A Critical Assessment, eds. John B. Gates and Charles A. Johnson. Washington, DC: Congressional Quarterly*, 281-301.
- MacKenzie, D. L. (2001). Corrections and sentencing in the 21st century: Evidence-based corrections and sentencing. *The Prison Journal, 81*, 299-312.
- Miethe, T. D. (1987). Charging and plea bargaining practices under determinate sentencing: An investigation of the hydraulic displacement of discretion. *Journal of Criminal Law and Criminology, 155-176*.
- Miethe, T. D. & Moore, C.A. (1985). Socioeconomic Disparities under Determinate Sentencing Systems: A Comparison of Pre- and Post-Guideline Practices in Minnesota. *Criminology, 23*, 337-363.
- Moore, C. A. & Miethe, T. D. (1986). Regulated and Unregulated Sentencing Decisions: An Analysis of First-Year Practices under Minnesota's Felony Sentencing Guidelines. *Law and Society Review, 20*, 253-277.

- Moore, L. & Vanneman, R. (2003). Context Matters: Effects of the Proportion of Fundamentalists on Gender Attitudes. *Social Forces*, 82, 115-139.
- Morris, H. (1981). A paternalistic theory of punishment. *American Philosophical Quarterly*, 263-271.
- Moulds, E. F. (1978). Chivalry and Paternalism: Disparities of Treatment in the Criminal Justice System. *Western Political Quarterly*, 31, 416-430.
- Mustard, D. B. (2001). Racial, ethnic, and gender disparities in sentencing: Evidence from the U.S. Federal Courts. *Journal of Law and Economics*, 44, 285-314.
- Myers, M. A., & Talarico, S. M. (1987). *The social contexts of criminal sentencing*. New York, NY: Springer-Verlag.
- Nagel, I. H. (1990). Structuring sentencing discretion: The new federal sentencing guidelines. *The Journal of Criminal Law & Criminology*, 80, 883-943.
- Nagel, I., & Hagan, J. (1983). Gender and crime: Offense patterns and criminal court sanctions. In *Crime and Justice: An Annual Review of Research*, vol. 4, eds. Michael Tonry and Norval Morris. Chicago: University of Chicago Press.
- Nagel, I. & Johnson, B. L. (1994). The Role of Gender in a Structured Sentencing System: Equal Treatment, Policy Choices, and the Sentencing of Female Offenders Under the United States Sentencing Guidelines. *Journal of Criminal Law and Criminology*, Summer, 181-221.
- Nagel, I. H., & Schulhofer, S. J. (1992). A tale of three cities: An empirical study of charging and bargaining practices under the federal sentencing guidelines. *Southern California Law Review*, 66, 501-566.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math= male, me= female, therefore math≠ me. *Journal of personality and social psychology*, 83, 44-59.
- Packer, H. (1968). *The limits of the criminal sanction*. Stanford University Press.
- Paternoster, R. (2011). Racial disparity under the federal sentencing guidelines pre- and post-Booker. *Criminology & Public Policy*, 10, 1063-1072.
- Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the Correct Statistical Test for the Equality of Regression Coefficients. *Criminology*, 36, 859-866.
- Peek, C. W., Lowe, G. L., & Williams, L. S. (1991). Gender and God's word: Another look at religious fundamentalism and sexism. *Social Forces*, 69, 1205-1221.

- Peterson, R. D., & Hagan, J. (1984). Changing conceptions of race: Towards an account of anomalous findings of sentencing research. *American Sociological Review*, 49, 56-70.
- Peterson, R. D., Krivo, L. J., & Hagan, J. (2006). *The many colors of crime: Inequalities of race, ethnicity, and crime in America*. NYU Press.
- Pollak, O. (1950). *The Criminality of Women*. Philadelphia: University of Pennsylvania Press.
- Portillos, E. L. (1998). Latinos, gangs, and drugs. In C. R. Mann & M. S. Zatz (Eds.), *Images of color, images of crime*. Los Angeles: Roxbury.
- Quinney, R. (1977). *Class, state, and crime: On the theory and practice of criminal justice (p.58)*. D. McKay Company.
- Quinney, R. (1970). *The social reality of crime*. Transaction publishers.
- Raeder, M. S. (1993). Gender and Sentencing: Single Moms, Battered Women, and Other Sex-Based Anomalies in the Gender-Free World of the Federal Sentencing Guidelines. *Pepperdine Law Review*, 20, 905-990.
- Raeder, M. S. (2006). Gender-Related Issues in a Post-Booker Federal Guidelines World. *McGeorge Law Review* 37, 691-756.
- Rafter, N. H. (1990). *Partial Justice: Women, Prison, and Social Control*. New Brunswick, NJ: Transaction Publishers.
- Rache, C. (1975). The female offender as an object of criminological research. In A. Brodsky (Ed.), *The female offender*. Beverly Hills, CA: Sage.
- Raudenbush, S. & Bryk, A. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage.
- Rehavi, M. Mt, & Starr, S. (2012). Racial Disparity in Federal Criminal Charging and its Sentencing Consequences. Working Paper no. 12-002. University of Michigan Law and Economics, Empirical Legal Studies Center, Ann Arbor, Mich.
- Rice, T. & Coates, D. (1995). Gender role attitudes in the southern United States. *Gender and Society*, 9, 744-756.
- Richardson, R. J., & Vines, K. N. (1970). *The politics of federal courts: Lower courts in the United States*. Little, Brown.
- Richey-Mann, C. (1984). Race and sentencing of female felons: A field study. *International Journal of Women's Studies* 7, 160-72.

- Roberts, D. E. (1994). Foreword: The Meaning of Gender Equality in Criminal Law. *Journal of Criminal Law and Criminology*, 85, 1-14.
- 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, 318-339.
- Rome, D. M. (1998). Murderers, rapists, and drug addicts. In C. R. Mann & M. S. Zatz (Eds.), *Images of color, images of crime*. Los Angeles: Roxbury.
- Rothman, D. J. (1971). *The discovery of the asylum*. Transaction Publishers.
- Rothman, D. J. (1983). Sentencing reforms in historical perspective. *Crime & Delinquency*, 29, 631-647.
- Rubin, A. B., & Bartell, L. B. (1989). *Law clerk handbook: a handbook for law clerks to Federal judges* (Vol. 89, No. 2). Federal Judicial Center.
- Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of social issues*, 57, 743-762.
- Rusche, G., & Kirchheimer, O. (1939). *Punishment and social structure*. New York: Russell & Russell.
- Sampson, R. J. (2012). *Great American city: Chicago and the enduring neighborhood effect*. University of Chicago Press.
- Sampson, R. J. & Laub, J. (1993). Structural variations in juvenile court processing: Inequality, the underclass, and social control. *Law and Society Review*, 27, 285-312.
- Savelsberg, J. (1992). Law that does not fit society: Sentencing guidelines as a neoclassical reaction to the dilemmas of substantivized law. *American Journal of Sociology*, 97, 1346-1381.
- Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, 8, 3-15.
- Schanzenbach, M. M. (2005). Racial and Gender Disparities in Prison Sentences: The Effect of District-Level Judicial Demographics. *Journal of Empirical Legal Studies*. 57, 57-92.
- Schanzenbach, M. M., & Tiller, E. H. (2007). Strategic judging under the US sentencing guidelines: Positive political theory and evidence. *Journal of Law, Economics, & Organization*, 23, 24-56.
- Schulhofer, S. J. (1992). Plea bargaining as disaster. *Yale Law Journal*, 1979-2009.

- Schur, E. M. (1984). *Labeling women deviant: Gender, stigma, and social control*. New York: McGraw-Hill.
- Shively, W. P. (1980). The nature of party identification: A review of recent developments. In J. C. Pierce & J. L. Sullivan (Eds.), *The electorate reconsidered* (pp.219-236). Beverly Hills: Sage publications, Inc. Smith, 2004
- Spitzer, S. (1975). Punishment and social organization: A study of Durkheim's theory of evolution. *Law & Society Review*, 9, 613-637.
- Spohn, C. (1999). Gender and sentencing of drug offenders: Is chivalry dead? *Criminal Justice Policy Review*, 9, 365–399.
- Spohn, C. (2000). *Thirty years of sentencing reform: The quest for a racially neutral sentencing process*. Washington, DC: National Institute of Justice: Criminal Justice.
- Spohn, C. (2009). *How do Judges Decide? The Search for Fairness and Justice in Punishment*. Thousand Oaks, CA: Sage.
- Spohn, C. (2011). Unwarranted disparity in the wake of the Booker/Fanfan decision. *Criminology & Public Policy*, 10, 1119-1127.
- Belenko, S., & Spohn, C. (2014). *Drugs, Crime, and Justice*. Sage Publications.
- Spohn, C., & Beichner, D. (2000). Is preferential treatment of female offenders a thing of the past? A multi-site study of gender, race, and imprisonment. *Criminal Justice Policy Review*, 11, 149-184.
- Spohn, C., & Fornango, R. (2009). US attorneys and substantial assistance departures: testing for interprosecutor disparity. *Criminology*, 47, 813-846.
- Spohn, C., & Holleran, D. (2000). The imprisonment penalty paid by young, unemployed black and Hispanic male offenders. *Criminology*, 38, 281-306.
- Spohn, C. C., Kim, B., Belenko, S., & Brennan, P. K. (2014). The Direct and Indirect Effects of Offender Drug Use on Federal Sentencing Outcomes. *Journal of Quantitative Criminology*, 30, 549-576.
- Spohn, C., & Spears, J. (1996). The Effect of offender and victim characteristics on sexual assault case processing decisions. *Justice Quarterly*, 13, 401-431.
- Stacey, A. M., & Spohn, C. (2006). Gender and the social costs of sentencing: An analysis of sentences imposed on male and female offenders in three U.S. district courts. *Boalt Journal of Criminal Law*, 11, 43-76.

- Stark, R., Doyle, D. P., & Kent, L. (1980). Rediscovering moral communities: Church membership and crime. *Understanding crime: current theory and research*, 18, 43-52.
- Starr, S. B. (2014). Estimating gender disparities in federal criminal cases. *American Law and Economics Review*, 14, 1-33.
- Starr, S. B., & Rehavi, M. M. (2013). Mandatory sentencing and racial disparity: Assessing the role of prosecutors and the effects of Booker. *Yale Law Journal*, 123, 2-48.
- Steffensmeier, D. (1980). Assessing the impact of the women's movement on sex-based differences in the handling of adult criminal defendants. *Crime and Delinquency*, 23, 344-356.
- Steffensmeier, D., & Demuth, S. (2006). Does gender modify the effects of race-ethnicity on criminal sanctioning? Sentencing for male and female white, black, and Hispanic defendants. *Journal of Quantitative Criminology*, 22, 241-261.
- Steffensmeier, D., J. Kramer., and C. Streifel. (1993). Gender and imprisonment decisions. *Criminology*, 31, 411-446.
- Steffensmeier, D., Ulmer, J.T., & Kramer, J.H. (1998). The interaction of race, gender, and age in criminal sentencing: the punishment cost of being young, black, and male. *Criminology*, 36, 763-797.
- Stith, K. (2008). The Arc of the Pendulum: Judges, Prosecutors, and the Exercise of Discretion. *Yale Law Journal*, 117, 1420-1497.
- Stith, K., & Cabranes, J. A. (1998). *Fear of judging: Sentencing guidelines in the federal courts*. Chicago, IL: University of Chicago Press.
- Sudnow, D. (1965). Normal crimes: Sociological features of the penal code in a public defender office. *Social problems*, 12, 255-276.
- Sutton, J. R. (1987). Doing time: dynamics of imprisonment in the reformist state. *American Sociological Review*, 52, 612-630.
- Tonry, M. (1987). Prediction and classification: Legal and ethical issues. *Crime and Justice*, 367-413.
- Tonry, M. (1993). Sentencing commissions and their guidelines. *Crime and Justice*, 137-195.
- Tonry, M. (1995). *Malign neglect: Race, crime, and punishment in America*. New York: Oxford University Press.
- Tonry, M. (1996). *Sentencing matters*. New York, NY: Oxford University Press.

- Tonry, M. (2009). Explanations of American punishment policies A national history. *Punishment & Society, 11*, 377-394.
- Tyler, T. R., & Weber, R. (1982). Support for the death penalty; instrumental response to crime, or symbolic attitude? *Law and Society Review, 17*, 21-45.
- Ulmer, J. T. (1997). *Social worlds of sentencing: Court communities under sentencing guidelines*. Albany, NY: State University of New York Press.
- Ulmer, J. T. (2012). Recent developments and new directions in sentencing research. *Justice Quarterly, 29*, 1-40.
- Ulmer, J. T., Bader, C., & Gault, M. (2008). Do moral communities play a role in criminal sentencing? Evidence from Pennsylvania. *The Sociological Quarterly, 49*, 737-768.
- Ulmer, J. T., Eisenstein, J., & Johnson, B. D. (2010). Trial penalties in federal sentencing: extra-guidelines factors and district variation. *Justice Quarterly, 27*, 560-592.
- Ulmer, J. T., & Johnson, B.D. (2004). Sentencing in context: a multilevel analysis. *Criminology, 4*, 137-178.
- Ulmer, J. T., & Kramer, J. H. (1996). Court communities under guidelines: Dilemmas of formal rationality and sentencing disparity. *Criminology, 34*, 383-408.
- Ulmer, J. T., & Kramer, J. H. (1998). The use and transformation of formal decision making criteria: Sentencing guidelines, organizational contexts, and case processing strategies. *Social Problems, 45*, 248-267.
- 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*, 427-458.
- Ulmer, J. T., & Light, M. T. (2011). Disparity: Changes in federal sentencing after Booker and Gall? *Federal Sentencing Reporter, 23*, 333-341.
- Ulmer, J. T., Light, M. T., & Kramer, J. H. (2011a). The “liberation” of federal judges’ discretion in the wake of the Booker/Fanfan decision: Is there increased disparity and divergence between courts? *Justice Quarterly, 28*, 799-837.
- Ulmer, J. T., Light, M. T., & Kramer, J. H. (2011b). Racial disparity in the wake of the Booker/Fanfan decision: An alternative analysis to the USSC’s 2010 report. *Criminology & Public Policy, 10*, 1077-1118.
- U.S. Census Bureau. (2000). Region and Country or Area of Birth of the Foreign-Born Population: 2000. Retrieved from http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_00_SF3_QTP15&prodType=table.

- U.S. Census Bureau (2011). Race data. Retrieved from <http://www.census.gov/population/www/socdem/race/racefactcb.html>.
- U.S. Sentencing Commission. (1991a). The Federal sentencing guidelines: A report on the operation of the guidelines system and short-term impacts on disparity in sentencing, use of incarceration, and prosecutorial discretion and plea bargaining (Vol. 2). Washington, DC: United States Sentencing Commission.
- U.S. Sentencing Commission. (1991b). Mandatory Minimum Penalties in the Federal Criminal Justice System. Washington, DC: United States Sentencing Commission.
- U.S. Sentencing Commission. (2003). Downward departures from the federal sentencing guidelines. Washington, DC: United States Sentencing Commission.
- U.S. Sentencing Commission. (2004). Fifteen years of guidelines sentencing: An assessment of how well the federal criminal justice system is achieving the goals of sentencing reform. Washington, DC: United States Sentencing Commission.
- U.S. Sentencing Commission. (2006). Final report on the impact of *United States v. Booker* on federal sentencing. Washington, DC: United States
- U.S. Sentencing Commission. (2008). Sourcebook of federal sentencing statistics. Washington, DC: United States Sentencing Commission.
- U.S. Sentencing Commission. (2009). Overview of federal criminal cases fiscal year 2008. Washington, DC: United States Sentencing Commission.
- U.S. Sentencing Commission. (2010). Demographic differences in federal Sentencing practices: An update of the Booker report's multivariate regression analysis. Washington, DC: United States Sentencing Commission.
- U.S. Sentencing Commission. (2012). *Report on the continuing impact of United States v. Booker on federal sentencing*. Washington, DC: Author.
- Vieraitis, L. M., Britto, S., & Kovandzic, T. V. (2007). The Impact of Women's Status and Gender Inequality on Female Homicide Victimization Rates Evidence From US Counties. *Feminist Criminology*, 2, 57-73.
- Von Hirsch, A. (1976). *Doing justice: The choice of punishments*. New York: Hill and Wang
- Walker, S. (1993). *Taming the system: The control of discretion in criminal justice, 1950-1990*. New York: Oxford University Press.
- Wang, X., & Mears, D. P. (2010a). A multilevel test for minority threat effects on sentencing. *Journal of Quantitative Criminology*, 26, 191-215.

- Wang, X., & Mears, D. P. (2010b). Examining the direct and interactive effects of changes in racial and ethnic threat on sentencing decisions. *Journal of Research in Crime and Delinquency*, 47, 522-557.
- Ward, G., Farrell, A., & Rousseau, D. (2009). Race effects of representation among federal court workers: Does black workforce representation reduce sentencing disparities? *Annals of the American Academy of Political and Social Science*, 623, 121-133.
- Weidner, R. R., & Frase, R. S. (2003). Legal and extralegal determinants of inter-county differences in prison use. *Criminal Justice Policy Review*, 14, 377-400.
- Weidner, R.R., Frase, R., & Pardoe, I., (2004). Explaining Sentence Severity in Large Urban Counties: A Multilevel Analysis of Contextual and Case-Level Factors, *Prison Journal*. 184, 184-207.
- Welter, B. (1978). She Hath Done What She Could: Protestant Women's Missionary Careers in Nineteenth-Century America. *American Quarterly*, 30, 624-638.
- Wheeler, S., Weisburd, D., & Bode, N. (1982). Sentencing the white collar offender: Rhetoric and reality. *American Sociological Review*, 47, 641-659.
- Wilkins Jr, W. W., & Steer, J. R. (1989). Relevant Conduct: The Cornerstone of the Federal sentencing Guidelines. *South Carolina Law Review*, 41, 495-531.
- Williams, J. E., & Best, D. L. (1982). *Measuring sex stereotypes: A thirty-nation study*. Sage Publications, Inc.
- Wilson, J. Q., (1975). *Thinking about crime*. New York: Basic.
- Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. University of Chicago Press.
- Wolgast, E. H. (1980). *Equality and the Rights of Women* (pp. 22-23). Ithaca: Cornell University Press.
- Woodberry, R. D., & Smith, C. S. (1998). Fundamentalism et Al.: Conservative Protestants in America. *Annual Review of Sociology*, 24, 25-56.
- Wooldredge, J. D. (1998). Analytical rigor in studies of disparities in criminal case processing. *Journal of Quantitative Criminology*, 14, 155-179.
- Wooldredge, J. (2010). Judges' unequal contributions to extra-legal disparities in imprisonment. *Criminology*, 48, 539-567.
- Wooldridge, J. M. (2009). *Introductory econometrics: a modern approach (4th)*, South-Western Cengage Learning.

- Wu, J., & Spohn, C. (2008). Interdistrict disparity in sentencing in three US district courts. *Crime & Delinquency*, 56, 290-322.
- Xie, M., Lauristen, J. L., & Heimer, K. (2012). Intimate partner violence in U.S. Metropolitan areas: The contextual influences of police and social services. *Criminology*, 50, 961-992.
- Yellen, D. N. (2000). Is Relevant Conduct Relevant-Reconsidering the Guidelines Approach to Real Offense Sentencing. *Saint Louis University Law Journal*, 44, 409-423.
- Young, V. (1986). Gender expectations and their impact on black female offenders and victims. *Justice Quarterly*, 3, 305-326.
- Zalman, M. (1977). Rise and fall of the Indeterminate Sentence, *Wayne Law Review*, 24, 45-79.