The Relationship between Race, Ethnicity, and Sentencing Outcomes: A Meta-Analysis of Sentencing Research

ABSTRACT

Statement of Purpose: A tremendous body of research has accumulated on the topic of racial and ethnic discrimination in sentencing. These studies have produced seemingly divergent findings. The purpose of this research is to conduct an objective, comprehensive, and systematic review of the literature regarding the relationship between race/ethnicity and sentencing outcomes using quantitative methods (i.e., meta-analysis), which remedy many of the shortcomings inherent in the extant qualitative (narrative) reviews. Further, this research goes beyond simply addressing the question of whether there is unwarranted racial/ethnic sentencing disparity, but also addresses the question of *why* this body of research produces such inconsistent findings.

Methods: This research employed meta-analysis to summarize and analyze the variability in research findings. This meta-analysis established explicit, pre-set eligibility criteria that governed inclusion decisions in this review. Hundreds of studies are retrieved; each was closely scrutinized to determine eligibility status. From each study, observed differences in sentencing outcomes by race/ethnicity, independent of defendant's criminal history and seriousness of the current offense, were transformed into effect sizes. We also coded features of each study's sample, methodology, type of sentencing outcome, and sentencing context.

Data Analysis: The data set was analyzed using meta-analytic analogues to traditional analysis of variance, and multiple-regression.

Results: Eighty-five studies meeting our stated eligibility criteria were located. Analysis of these data reveal that, after taking into account defendant criminal history and current offense seriousness, African-Americans and Latinos were generally sentenced more harshly than whites. Differences in sentencing outcomes between these groups generally were statistically significant but statistically small (although not necessarily substantively small). Further, analyses indicate that larger estimates of unwarranted sentencing disparity were found in studies that examined drug offenses, imprisonment or discretionary sentencing decisions, and in recent analyses of Federal court data. Smaller estimates of unwarranted sentencing disparity were found in analyses that employed more control variables (especially those that controlled for defendant SES), utilized precise measures of key variables, or examined sentencing outcomes relating to length of incarcerative sentence. Additionally, there was some evidence to suggest that structured sentencing mechanisms, such as sentencing guidelines, were associated with smaller unwarranted sentencing disparities. The limited available research contrasting sentencing patterns of whites to those of Asians or Native Americans does not generally reveal significant differences between these groups.

Conclusions: Overall, these findings call into question the so-called "no discrimination thesis." These findings suggest that policy-makers need to re-evaluate sentencing practices, especially in regards to drug offenses and the decision to incarcerate.

The Relationship between Race, Ethnicity, and Sentencing Outcomes:

A Meta-Analysis of Sentencing Research

RESEARCH SUMMARY

FINAL REPORT

Submitted to the National Institute of Justice

Ojmarrh Mitchell, Ph.D.

Department of Criminal Justice

University of Nevada, Las Vegas

Doris L. MacKenzie, Ph.D.

Department of Criminology and Criminal Justice

University of Maryland

This project was supported by Grant No.: 2002-IJ-CX-0020 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice

The Relationship between Race, Ethnicity, and Sentencing Outcomes: A Meta-Analysis of Sentencing Research

RESEARCH SUMMARY

The issue of racial and ethnic disparity in criminal sentencing has been one of the longest standing research topics in all of criminology. At least 70 years of empirical research has focused on this issue without a clear consensus emerging. Over that period, a tremendous body of research has accumulated on this topic. Some studies have found that racial/ethnic minorities are sentenced more harshly than whites even after legally relevant factors, such as offense seriousness and prior criminal history, are taken into consideration. Conversely, a few studies have reached the opposite conclusion–racial minorities are treated *more leniently* than whites, while still other research has found no differences in sentencing outcomes by race/ethnicity of the defendant.

The current research reports the results from a quantitative (i.e., meta-analytic) synthesis of empirical research assessing the influence of race/ethnicity on non-capital sentencing decisions in U.S. criminal courts. Eighty-five studies meeting our eligibility criteria were located. From each of these studies, the magnitude and direction of observed racial/ethnic disparities were calculated (via effect sizes). That is, from each study we measured the actual size and direction (i.e., which racial/ethnic group was disadvantaged) of any observed differences in sentencing outcomes by race/ethnicity. Analyses of these data not only determine whether there is racial/ethnic disparity disadvantaging minorities, but also estimate the magnitude of such disparity. Perhaps most importantly, the results of these analyses go beyond addressing the simple question

of whether unwarranted racial disparity exists and attempts to address the question of *why* studies of racial disparity in sentences often produce inconsistent findings. We believe that the inconsistent findings exhibited in this body of research can be explained by taking into account the varying features of each study and characteristics of the sentencing jurisdiction. Specifically, we believe that variations in methodology, sample, and context are *systematically* related to the results produced by the empirical research.

Several authors have attempted to review this voluminous and diverse body of research using traditional qualitative narrative literature review techniques. In many regards, these reviews are insightful and invaluable; however, it is our contention that these reviews are of limited utility because of the qualitative, narrative methodology employed in each. Perhaps the most important weakness typically exhibited by these narrative reviews is that they overemphasize statistical significance of individual findings. Existing narrative reviews most often use a "vote-counting" methodology that simply determines the proportion of studies finding a statistically significant effect, regardless of the magnitude of this effect. As a result, narrative reviews provide answers to conceptually flawed questions, such as: What proportion of studies show a statistically significant effect of race/ethnicity on sentencing outcomes? This is a flawed question, as in the extreme, this question is nothing more than asking: How many studies used large samples?

The existing narrative reviews are also plagued by a number of other shortcomings. For instance, traditional narrative reviews of the literature most often are not comprehensive. Most of the existing reviews of the race and sentencing literature examined only published studies. This is a significant shortcoming because previous

research in other areas of study have consistently demonstrated that studies reporting statistically significant results are more likely to be published. This leads to what is often referred to as "publication bias." Traditional narrative reviews also, at least implicitly, give equal weight to studies of varying methodological rigor. Even the most casual perusal of sentencing research reveals that there is a significant amount of variation in the methodological rigor of this body of research. Shouldn't studies with greater levels of methodologically rigor and/or larger samples be more heavily relied upon in drawing conclusions? Moreover, the existing reviews of sentencing research often include studies based on the same data or very similar data. This practice leads to double or triple counting what, in essence, is the same study.

Most narrative reviews tend to gloss over these issues; however, these shortcomings can lead to the conclusions of narrative reviews being inconsistent with the data. An article in *Science* (Mann 1994) provides a powerful illustration of this possibility. Mann compared the conclusions of meta-analytic reviews to those of narrative reviews in five areas of research and found that narrative reviews underestimated the presence and the strength of effects in each of the five research areas.

Given the shortcomings of existing narrative reviews of the race/ethnicity literature, we assert that these narrative syntheses have not utilized the *information* contained in the empirical research in a manner that maximizes *knowledge* regarding the relationship between race/ethnicity and sentencing decisions. What's needed is a method that objectively, systematically, and comprehensively reviews the literature regarding the relationship between race and sentencing outcomes and that can provide answers to the

substantive issues at hand, instead of focusing on statistical significance of individual findings.

In recent years, the method of "meta-analysis" has become increasingly popular in various fields including medicine, psychology, and crime prevention research. Meta-analysis has become the preferred method for synthesizing quantitative research because it solves many of the problems of narrative literature reviews. Instead of focusing on statistical significance, meta-analysis focuses on both the observed *direction* and *magnitude* of a relationship by using numerical effect size estimates. Meta-analysis gives more weight to larger studies, and provides a means of controlling for the methodological rigor of studies. In short, standard meta-analysis procedures yield focused, comprehensive, and systematic reviews of a body of research.

This study departs from earlier narrative reviews of the race and sentencing literature in several important ways. First, this research utilizes meta-analytic techniques that systematically and comprehensively review all available research (published and unpublished) pertaining to the influence of race and ethnicity on sentencing outcomes, meeting specific eligibility criteria. Second, this study reviews research on both race and ethnicity. Prior reviews of the literature tend to focus on contrasting sentencing outcomes between African-Americans and whites; whereas, the current study also contrasts sentencing outcomes between whites and Latinos, Asians, and Native Americans. Moreover, the meta-analytic procedure utilized in the current research enables us to address the question of *why* studies of racial disparity in sentences often produce inconsistent findings.

The specific aim of this study is to conduct a systematic and comprehensive meta-analysis that examines: 1) whether race and ethnicity are related to sentencing outcomes, independent of offense seriousness and defendant criminal history; 2) in which contexts are racial/ethnic bias most likely to occur (e.g., Southern jurisdictions, jurisdictions without structured sentencing); and, 3) whether study findings are systematically related to variations in research methodology, sample, and context.

The present meta-analytic review conducted a search for published and unpublished studies that assessed the influence of race or ethnicity on sentencing outcomes. Bibliographic databases (i.e., PsychLit, MedLine, NCJRS, Criminal Justice Abstracts, Sociological Abstracts, Social Science Citation Index, SocioFile, Conference Papers Index, and UnCover), reference lists from literature reviews, conference proceedings, hand searches of select relevant journals, and dissertations via Dissertation Abstracts were searched for potentially eligible evaluations. Potentially eligible evaluations were retrieved and closely scrutinized to determine eligibility for this review. Additionally, we contacted each state's sentencing body to determine whether these organizations have internally evaluated their jurisdiction's sentencing practices in regards to unwarranted racial/ethnic sentencing disparity.

The eligibility criteria for inclusion in the present research were that each study had to: 1) be conducted using cases sentenced in the United States; 2) examine sentencing outcomes in criminal courts (i.e., juvenile court adjudication/sentencing outcomes are excluded); 3) incorporate *simultaneous* controls for both offense seriousness and criminal history; 4) measure the direct influence of race/ethnicity on sentencing outcomes; 5)

examine sentencing outcomes unrelated to death penalty decisions; and, 6) all research must be made available through yearend 2002.

This meta-analysis focuses on five types of sentencing outcomes: 1) imprisonment decisions, 2) length of incarcerative sentence, 3) simultaneous examinations of imprisonment and sentence length decisions, 4) discretionary lenience, and 5) discretionary punitiveness. Imprisonment and length of incarcerative sentence decisions are self-explanatory. The third type of sentencing outcome, what we refer to as simultaneous examinations of imprisonment and sentence decisions, typically analyze sentencing decisions using ordinal scales of measure with non-incarcerative sentences (i.e., probation) being the least severe sentence, short-term incarcerative sentences being the next level of severity, and longer-term incarcerative sentences being progressively more severe on the ordinal scale. Discretionary lenience refers to sentencing outcomes where judges sanction defendants in some manner more lenient than ordinary (e.g., downward departures from guidelines, stays of sentence). Conversely, discretionary punitiveness refers to sentencing outcomes where judges sanction defendants more harshly than ordinary (e.g., upward departures, enhanced sentencing provisions for eligible repeat offenders).

From each independent sentencing context, an odds-ratio effect size was calculated for each minority-white comparison of sentencing outcomes. We chose the odds-ratio effect size because the majority of the minority-white comparisons were conducted with dichotomous measures of sentencing outcomes (e.g., imprisonment vs. probation). Effect sizes were coded in manner such that *larger effect sizes indicate minorities were punished more harshly than whites.* Specifically, odds-ratios greater than

1 indicated that the minority group of interest were sentenced more harshly than whites, independent of defendant criminal history and current offense seriousness, with larger odds ratios denoting larger minority-white sentencing disparities; whereas odds-ratios less than 1 indicated that whites were sentenced more harshly than minorities.

Key features of each study's methodology (e.g., number and nature of control variables, how offense seriousness and criminal history were measured), sample (type of offenses analyzed, age of sample, gender distribution of sample), type of sentencing outcome, publication status (published vs. unpublished), and sentencing context (e.g., region, use of guidelines, time period of data) were coded. These coded variables are used as independent variables to predict variation in the magnitude of observed sentencing disparities between races/ethnicities.

A total of 184 studies meeting the stated eligibility criteria were located. Eighty of these studies were excluded from the analyses because they were determined to be statistically dependent (i.e., analyzed the same data set as another study included in this synthesis) and 19 studies were excluded because they were uncodeable (usually because some vital piece of statistical information was not reported). Finally, nine studies analyzed the same data as another study but analyzed a different outcome, these nine studies were merged with their corresponding studies–leaving 76 eligible and statistically independent studies (these 76 studies encompass the results of 85 studies). Approximately half of the studies included in this research were published as journal articles (49%), another 14% of studies were published as books or book chapters, and a considerable proportion of studies were unpublished (37%).

The descriptive analysis of the studies included in this meta-analysis reveals several important points. First, the current research appears to be the most comprehensive review of this research, as this meta-analysis includes the results from 85 published and unpublished studies. Further, this meta-analysis also appears to be comprehensive in regards to regional representation, as analyses of unwarranted racial disparity from 29 states, the Federal courts, and Washington D.C. were included. Most of the included studies analyzed data from the 1970s or 1980s (76%), whereas 17% of studies utilized more recent data.

Second, it is evident that the methodological rigor in this body of research has improved markedly over the past several decades. Analyses of unwarranted racial/ethnic disparity have included an increasing number of control variables. Most commonly, these controls measure defendant socioeconomic status (SES), type of defense counsel, and method of case disposition—all of which have been found to be related to race/ethnicity and severity of sentencing outcomes. The precision of measurement of control variables utilized has also improved; for instance, whereas many of the early studies employed crude dichotomous measures of important variables, such as offender criminal history, considerably fewer recent studies employ variables measured in this manner.

Third, most sentencing research focuses on comparing sentencing outcomes of African-Americans to whites. In fact, 95% of coded studies included contrasts between African-Americans and whites, and 66% of all effect sizes contrasted sentencing outcomes of African-Americans to those of whites. However, an increasing number of studies are including empirical assessments of Latino sentencing outcomes. In all, 25%

of coded contrasts compared sentencing outcomes of Latinos to whites and almost all of this research was published since 1980. Little research concerns sentencing practices of Native Americans or Asians as only 5% and 4% of coded contrasts, respectively, compared sentencing outcomes between these minority groups and whites.

The effect size analyses revealed that even after taking into consideration current offense seriousness and defendant prior criminal conduct, African-Americans and Latinos were sentenced more harshly, on average, than whites. In regards to contrasts between African-Americans and whites, the magnitude of sentencing disparity was highly variable, and varied by type of court (State or Federal) in particular. The mean overall odds-ratio for State sentencing contexts was 1.28, which is statistically significant, whereas the mean overall effect size for Federal sentencing contexts was 1.15. It is very important to note, however, that more recent analyses of Federal data yield considerably greater indications of unwarranted racial disparity; specifically, in analyses of Federal data collected since 1980, the mean odds ratio effect size was 1.58 compared to a mean of 1.02 for analyses conducted with data before 1980. Translating these mean odds ratios in percentages, while assuming a 50% rate of punishment for whites, suggests that African-Americans in State courts would be punished at a rate of 56% and African-Americans in Federal courts would be punished at a rate of 53%; restricting attention to only more recent analyses of Federal data would increase the percentage of African-Americans expected to be punished to 61%.

The present research also found that the magnitude of the unwarranted sentencing disparity disadvantaging African-Americans varied by several other factors. First, estimates of unwarranted racial disparity varied by type of sentencing outcome.

Unwarranted sentencing disparity disadvantaging African-Americans was largest when imprisonment and discretionary sentencing decisions were scrutinized and smaller when decisions pertaining to length of incarcerative sentence were assessed. A second important source of systemic variation was type of offense. Specifically, sentencing disparity was largest in cases that examined sentences for drug offenses and smallest in cases involving property crimes. Third, several methodological features were important moderators of effect size. Those studies that utilized only a single dichotomous measure of criminal history generated considerably larger estimates of unwarranted disparity than analyses using more precise measures of this important variable. Likewise, analyses using imprecise measures of offense severity also produced larger estimates of unwarranted racial disparity. Smaller estimates of unwarranted racial disparity in sentencing were also found in analyses employing more control variables for factors known to be related to both severity of sentence and race, especially SES of defendant. Moreover, unpublished studies were associated with smaller estimates of unwarranted racial disparity than published studies-suggesting that narrative reviews which focus on published research utilized a biased sub-sample of all available research. Another interesting finding is that analyses of highly aggregated data (e.g., analyses of data pooled from all jurisdictions within one state) produced systematically smaller estimates of unwarranted sentencing disparity than analyses conducted at a lower (smaller) levels of aggregation. This finding strongly suggests that aggregation bias affects estimates of unwarranted racial/ethnic disparity in analyses using higher levels of aggregation. Yet, even after taking into account these various methodological features, our analyses indicate that unwarranted disparity disadvantaging African-Americans persisted.

It is also important to note which variables failed to exhibit meaningful associations with effect size. Perhaps most salient, is the weak negative association between effect size and presence of structure sentencing in analyses of State data. At the bivariate level, jurisdictions employing structured sentencing displayed smaller average levels of unwarranted racial disparity than jurisdictions without such mechanisms, and this association was marginally statistically significant. This difference, however, was not statistically significant once other factors were taken into consideration. Subsequent multivariate analyses found that while the presence of structured sentencing continued to have a modest negative relationship to effect size, this relationship was not statistically significant. Similarly, at the bivariate level, Southern jurisdictions displayed larger estimates of unwarranted disparity than other regions of the U.S., but this relationship was not statistically significant after taking into account methodological differences between studies.

While a considerably smaller number of studies analyzed contrasts between Latinos and whites, analyses of these contrasts found that Latinos in both State and Federal courts generally were sentenced more harshly than whites. Similar to the findings of African-American/white contrasts, the influence of ethnicity (i.e., being Latino in comparison to being white) was highly variable. The mean odds-ratio effect size in these Latino/white comparisons (1.18) was statistically small but statistically significant. Also similar to the findings of African-American effect sizes, unwarranted Latino/white sentencing disparity was greatest when imprisonment and discretionary sentencing outcomes were examined and in offenses involving drugs. In contrast to the analyses of African-American/white contrasts, methodological features of the analyses

were not as strongly related to effect size. In fact only two methodological characteristics, the use of selection bias corrections and number of variables measuring criminal history, were statistically related to effect size.

In regards to sentencing outcomes of Asians and Native Americans, few contrasts were available. The findings from the few available studies, however, indicate that sentencing disparity between these racial groups and whites were statistically small. The average difference in sentencing outcomes between Native Americans and whites was very small and not statistically significant. The difference between Asians and whites was not statistically or substantively significant in State courts; however, the difference between these groups was statistically significant but small in analyses of Federal court data. The small number of available effect sizes, however, makes drawing firm conclusions for these analyses tenuous.

These findings undermine the so-called "no discrimination thesis" which contends that once adequate controls for other factors, especially legal factors (i.e., criminal history and severity of current offense), are controlled unwarranted sentencing disparity disappears. Our analyses indicate that even after taking legal factors into account, Latinos and African-Americans were sentenced more harshly than whites on average. The observed differences between whites and these minority groups generally were statistically small suggesting that discrimination is not the primary cause of the overrepresentation of minorities in U.S. prisons. The extant literature also shows, however, considerable unwarranted racial and ethnic disparities in analyses involving drug offenses, imprisonment decisions, and recently collected Federal data.

Finally it needs to be realized that the preceding estimates of unwarranted racial/ethnic disparity were calculated *per sentencing episode*. The *cumulative disadvantage* endured by minorities may be considerably greater when multiple sentencing episodes are considered. That is, given the strong relationship between prior criminal history and sentencing outcomes, small disadvantages suffered in the past may have substantial effects in subsequent sentencing episodes. It is an undeniably criminological fact that one of the best predictors of future criminality is prior criminality. The implication of this finding is that many offenders will cycle through the criminal justice system repeatedly. Over time, as offenders repeatedly cycle through the criminal justice system, the small disadvantages suffered in each sentencing episode grow and may become *substantial* disadvantages.

These findings have implications policy-makers. First and foremost, the present research indicates that policy-makers need to (re-)examine the racial and ethnic neutrality of the sentencing policies and practices both generally and especially in regards to certain specific types of sentencing decisions. The persistent finding of differences in sentencing outcomes between minorities and whites suggests that policy-makers' efforts to achieve racial/ethnic neutrality have not been completely successful in eliminating such disparities. Yet, this research does provide some evidence that structured sentencing mechanisms at the State level are associated with smaller unwarranted sentencing disparities—signifying that these interventions have been at least marginally successful in this regard. Furthermore, the relatively larger sentencing disparities evident in imprisonment decisions and drug offenses suggests that policy-makers need to re-evaluate, and potentially alter, sentencing policies in these arenas.

The Relationship between Race, Ethnicity, and Sentencing Outcomes:

A Meta-Analysis of Sentencing Research

FINAL REPORT

Submitted to the National Institute of Justice

Ojmarrh Mitchell, Ph.D.

Department of Criminal Justice

University of Nevada, Las Vegas

Doris L. MacKenzie, Ph.D.

Department of Criminology and Criminal Justice

University of Maryland

This project was supported by Grant No.: 2002-IJ-CX-0020 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

The issue of racial and ethnic disparity in sentencing and imprisonment long has been a central concern of criminal justice research. As long ago as the 1920s, Sellin (1928) noticed stark disparities in incarceration rates between whites and African-Americans. His research revealed that, in comparison to whites, African-Americans were incarcerated at a rate of nearly 14 to 1. Recent comparisons of imprisonment rates continue to reveal racial and ethnic disparities; for example, at yearend 2001, Hispanic males were imprisoned at a rate two and half times greater than that of white non-Hispanic males, while African-American males were imprisoned at a rate of over seven and half times greater than non-Hispanic white males (Harrison and Beck 2002: 12).

Racial/ethnic discrimination is only one of several explanations that can be used to account for these disparities in imprisonment rates; as such, it is important to draw a distinction between *disparity* and *discrimination*. Racial/ethnic disparity simply indicates the existence of differences between racial or ethnic groups on some variable of interest. Disparities in imprisonment rates can arise due to many *legitimate* reasons; for example, differences in criminal history or seriousness of offense between racial/ethnic groups may produce racial/ethnic disparities in rates of imprisonment and length of sentence. In contrast, racial/ethnic discrimination refers to differences on some variable of interest between racial/ethnic groups that stem from biased treatment of some particular group. Further, this differential treatment is independent of an individual's past or present behavior.

It follows from this discussion of the distinction between racial disparity and discrimination that there are two main explanations for racial and ethnic differences in imprisonment. Some scholars contend that group differences in imprisonment rates can be explained by differences in legally relevant factors such as offense seriousness, prior criminal record, and other legally legitimate factors (Wilbanks 1987). That is, from this perspective African-Americans and Hispanics commit more serious offenses and have more serious prior criminal records than whites, which lead these groups to have higher rates of imprisonment. Other scholars contend that these differences in imprisonment rates are due to differences in legally relevant factors *and racial discrimination*. From this point of view, criminal justice decision-makers, faced with two similarly situated offenders, one white and the other a racial or ethnic minority, will view the minority's offense as being more threatening, or view the minority offender as less amenable to treatment and therefore punish the minority offender more harshly than the white offender (Steffensmeier, Ulmer, and Kramer 1998).

A tremendous body of research has accumulated on the topic of racial discrimination in sentencing without reaching a clear consensus. Some studies have found that racial/ethnic minorities are sentenced more harshly than whites even after legally relevant factors, such as offense seriousness and prior criminal history, are taken into consideration (Albonetti 1997; Bushway and Piehl 2001; Crawford, Chiricos, and Kleck 1998; Kramer and Steffensmeier 1993; Petersilia 1983). Conversely, a few studies have reached the opposite conclusion–racial minorities are treated *more leniently* than whites (Bernstein, Kelley, and Doyle 1977; Feimer, Pommersheim, and Wise 1990; Myers and Talarico 1986; Peterson and Hagan 1984), while still other research has found

no differences in sentencing outcomes by race/ethnicity of the defendant (Alvarez 1996; Engen and Gainey 2000; Klein, Petersilia, and Turner 1990; Lotz and Hewitt 1977). The findings from this body of research are further complicated by a growing body of primarily recent research that indicates race influences sentencing outcomes only in certain contexts, such as in certain geographical areas, types of cases (e.g., drug offenses), or in interaction with other factors (e.g., age and sex, historical period) (Crawford 2000; Nelson 1992; Spohn and Holleran 2000; Steffensmeier, Ulmer, and Kramer 1998; Zatz 1987).

Several authors have attempted to review this voluminous and diverse body of research using traditional qualitative narrative literature review techniques (e.g., Chiricos and Crawford 1995; Kleck 1981; Spohn 2000a). In many regards, these reviews are insightful and invaluable; however, it is our contention that these reviews are of limited utility because of the qualitative, narrative methodology employed in each. Perhaps the most important weakness typically exhibited by these narrative reviews is that they overemphasize statistical significance of individual findings. This weakness is particularly relevant in studies of sentencing research, as sample size in this body of research tends to be bipolar. At one extreme, contemporary studies tend to have very large samples. For example, in Spohn's (2000) review of recent sentencing studies (i.e., studies based on data collected since 1980) the average sample size in the included studies was over 30,000. The large samples used in these recent studies result in small differences between racial/ethnic groups being statistically significant. At the other extreme, studies conducted prior to the early 1980s tend to have considerably smaller samples. For example, in Hagan and Bumiller's (1983) review of earlier studies most

studies had sample sizes less than 1,000 with several studies having sample sizes of a few hundred cases. In these smaller studies, large racial sentencing disparities may not attain statistical significance due in part to low statistical power.

This focus on *statistical* significance instead of *substantive* significance shifts attention away from the most salient research questions towards conceptually flawed ones. That is, narrative reviews do not provide answers to important questions such as: How much more likely are defendants of one racial/ethnic group to receive an incarcerative sentence in comparison to whites? Or, how much longer are sentences imposed on defendants of a specific racial/ethnic group in comparison to whites? Instead of addressing these most fundamental and important research questions, existing narrative reviews must often use a "vote-counting" method that simply determines the proportion of studies finding a statistically significant effect, regardless of the magnitude of this effect. As a result, narrative reviews provide answers to conceptually flawed questions, such as: What proportion of studies show a statistically significant effect of race/ethnicity on sentencing outcomes? In the extreme, this question is nothing more than asking: How many studies used large samples?

Furthermore, extant narrative reviews of the literature most often are not comprehensive. Most of the reviews noted above examined only published studies. This is a significant shortcoming because previous research in other areas demonstrate that studies reporting statistically significant results are more likely to be published (Callaham, Wears, Barton, and Young 1998; Greenwald 1975; Smith 1980). Thus, these reviews utilize potentially non-representative, biased samples of studies, which in turn can lead to biased conclusions (i.e., "publication bias").

Other important shortcomings of traditional narrative reviews are that they give equal weight to studies of varying methodological rigor and often include several studies based on the same data. Even the most casual perusal of sentencing research reveals that there is a significant amount of variation in the methodological rigor of this body of research. For example, a number of studies use crude measures of race, offense severity, and prior criminal history, whereas other research uses more finely graded measures (Wooldredge 1998). Moreover, several recent reviews of sentencing research include studies based on the same or overlapping data sets. This practice leads to double or triple counting of what in essence is the same study. These narrative reviews tend to gloss over both of these issues, which could lead to significant distortion of the cumulative research findings.

These shortcomings may seem at first glance to be trivial, technical issues; however, there is evidence that these weaknesses can lead to a review's conclusions being inconsistent with the extant research. An article in *Science* (Mann 1994) provides a powerful illustration of this possibility. Mann compared the conclusions of meta-analytic reviews to those of narrative reviews in five areas of research, including delinquency prevention. Mann found that *narrative reviews underestimated the presence and the strength of effects in each of the five research areas*.

Another example of the susceptibility of narrative reviews to bias is given by Cooper and Rosenthal (1980). These authors randomly assigned researchers uninitiated with meta-analysis to two groups. The first group was instructed to review and synthesize the results for seven studies examining sex differences in task persistence using whatever method they would usually employ. The second group was given a

tutorial in meta-analysis then asked to review the same seven studies. Cooper and Rosenthal found that the meta-analytic group was nearly three times as likely to reject the null hypothesis of no difference between males and females in persistence as the narrative group, which in fact was the correct conclusion.

Given the substantial shortcomings of existing reviews of the race/ethnicity literature, it is our position that the field has yet to squeeze all of the available knowledge out of the existing empirical research. What is needed is a method that objectively, systematically, and comprehensively reviews the literature regarding the relationship between race and sentencing outcomes and that focuses on the substantive issue of the magnitude and direction of unwarranted racial disparity, instead of focusing on statistical significance.

In recent years, the method of "meta-analysis" has become increasingly popular in various fields including medicine, education, psychology, and crime prevention research. Meta-analysis remedies the shortcomings of traditional narrative reviews by using quantitative procedures to synthesize the findings from a collection of studies and describes the results from these studies using numerical "effect-size" estimates (Lipsey and Wilson 2001; Wang and Bushman 1999). By utilizing these numerical effect sizes, meta-analysis shifts the focus away from simply determining whether an effect is statistically significant to the direction and magnitude of that relationship, and thus provides a more meaningful indicator of the relationship. Standard meta-analytic practice requires the use of both published and unpublished research, and thus minimizes the problem of publication bias. Meta-analysis also provides a method of controlling for the various levels of methodological rigor inherent in reviews of a body of research, and

gives larger weight to studies with more precision (i.e., studies with smaller standard errors).

In sum, meta-analysis has become the preferred method for synthesizing empirical research because it solves many of the problems commonly encountered by narrative literature reviews. Instead of focusing on statistical significance, meta-analysis focuses on both the observed direction and magnitude of an effect. Standard meta-analytic practice requires systematic, focused, and comprehensive reviews. Meta-analysis gives more weight to larger studies, and provides a means of taking into account variability in methodological rigor.

This study departs from earlier reviews of the race and sentencing literature in several important ways. First, this research utilizes meta-analytic techniques that systematically and comprehensively review all available research (published and unpublished) pertaining to the influence of race and ethnicity on sentencing outcomes, meeting minimal eligibility criteria. Second, this study reviews research on both race and ethnicity. Prior reviews of the literature tend to focus solely or predominantly on contrasting sentencing outcomes between African-Americans and whites. This review also separately contrasts sentencing outcomes of racial and ethnic minorities with those of whites.

Perhaps most importantly, this review goes beyond the simple question of whether unwarranted racial disparity exists and attempts to address the question of *why* studies of racial disparity in sentences often produce inconsistent findings. We agree with Hagan and Bumiller's conclusion that: "The challenge is to explain why some studies find discrimination while others do not" (p. 31). We believe that the inconsistent

findings exhibited in this body of research can be explained by taking into account the varying features of each study and characteristics of the sentencing jurisdiction. Specifically, we believe that variations in methodology, sample, and context are *systematically* related to the results produced by the empirical research.

Research Questions

The specific aim of this study is to conduct a systematic and comprehensive meta-analysis that examines: 1) whether race and ethnicity are related to sentencing outcomes, independent of offense seriousness and defendant criminal history; 2) in which contexts are racial/ethnic bias most likely to occur (e.g., Southern jurisdictions, jurisdictions without structured sentencing); and, 3) whether the results from existing empirical studies are systematically related to methodological and sample variations.

Outline of Research

This report details the research questions, methodology, and analytic strategy for a meta-analytic review of the literature concerning differences in sentencing outcomes by race/ethnicity of the defendant in non-capital offenses. Chapter 2 reviews the findings of prior syntheses of this research and critiques the methodology of these prior syntheses. Chapter 3 details the research methodology, analytic strategy and limitations of this research. The results of the analyses are presented in Chapter 4. Finally, Chapter 5 discusses this study's findings and implications.

Definitional Issues

Above, we noted the general distinction between disparity and discrimination. In this section, we offer definitions of these terms specific to sentencing research. Following Blumstein and colleagues (Blumstein, Cohen, Martin, and Tonry 1983), *disparity* in sentencing "exists when 'like cases' with respect to case attributes—regardless of their legitimacy—are sentenced differently" (p. 72). Once again, disparities in sentencing outcomes can arise due to many *legitimate* reasons; for example, differences in criminal history or nature of the current offense between racial/ethnic groups may produce racial/ethnic disparities in sentencing outcomes. By contrast, Blumstein et al. note that *discrimination* in sentencing "exists when some case attribute that is objectionable (typically on moral or legal grounds) can be shown to be associated with sentence outcomes *after all other relevant variables are adequately controlled*" (p. 72, emphasis added).

Given this definition of discrimination, it is virtually impossible to statistically prove the existence of discrimination, as one rarely knows whether "all other relevant variables" have been included—not to mention adequately measured. This point is made clear by the arguments of Wilbanks (1987), who contends that the apparent "discrimination" revealed by some research on sentencing is artificial; a statistical illusion created by the correlation between race and omitted (unmeasured) variables that are also associated with sentencing outcomes. After reviewing the race and sentencing literature, Wilbanks concludes "[estimates of unwarranted racial disparity] may be the result of a race effect, but it may also stem from numerous other factors that were not controlled" (p. 109).

Given the limitations of statistical models to prove the existence of discrimination and the distinct possibility that at least some of the observed differences between minorities and whites are likely due specification error, we believe that the term "unwarranted disparity" is appropriate. Stated differently, it is difficult, if not impossible, to determine statistically whether the observed differences between minorities and whites ethnicities in sentencing outcomes are due to discrimination or reflects statistical misspecification. Thus, we use the term unwarranted disparity to emphasis this difficulty. Other terms could be used to describe this situation, for example others have suggested the use of "unexplained racial variation" as a more appropriate label (Wilbanks 1987). This label, however, implies that observed racial differences would disappear completely, if other unmeasured variables were included. In our view, this position assumes too much—whether other variables can explain away differences between minorities and whites remains to be seen.

It is also important to define the terms "race" and "ethnicity." *Race and ethnicity are social constructs.* In our society, race is most often defined in terms of skin color, whereas ethnicity refers to an individual's cultural and ancestral origin. That is, ethnicity refers to group variations in region of family origin, customs, language, diet, and religion. Further, within any race there may be numerous ethnic groups. For example, individuals classified as "whites" can be further classified as Irish, Italian, French, and so forth.

The terms race and ethnicity may at first glance seem to be objective and non-problematic; however, upon closer inspection flaws emerge in these definitions. The best evidence of race as a social construct is found by comparing racial definitions between societies. A recent article in the *Seattle Times* exemplifies this inconsistency. In this

article a Brazilian immigrant is quoted as remarking: "In this country if you are not quite white, then you are black, [but in Brazil] if you are not quite black, then you are white" (Fears 2003). The term Latino (or Hispanic) also complicates the distinction between race and ethnicity. Most frequently, this term is used as an ethnic category; however, this is problematic because Latinos are in fact comprised of a host of ethnic and racial groups including black and white Cubans, Puerto-Ricans, Brazilians, and so forth. These difficulties underscore the fact that these terms are arbitrary, social constructs that change over time and place. Yet, within time and place, these terms are reliable.

In this research, we use the term "race" to refer to socially accepted classifications of individuals based primarily on skin color. The term "ethnicity" refers primarily to differences in family original and culture. We use the term Latino to differentiate primarily Spanish-speaking ethnic groups tracing their ancestral origins to North, Central, and South America. Lastly, we use the term "minorities" to denote racial and ethnic minorities (i.e., people who are not non-Hispanic whites).

CHAPTER 2. PRIOR RESEARCH

The racial and ethnic disparity in sentencing outcomes is a long-standing issue in the fields of sociology and criminology. Yet after nearly 70 years of sustained empirical research on this topic, little consensus exists. Some studies, perhaps a majority, find that racial or ethnic minorities are punished more harshly than comparably culpable whites convicted of similar crimes, at least in some types of sentencing outcomes or types of offenses (Albonetti 1997; Bushway and Piehl 2001; Crawford, Chiricos, and Kleck 1998; Kramer and Steffensmeier 1993). In contrast, another sizeable portion of this research finds no differences in sentencing outcomes by race/ethnicity of the defendant (Chiricos and Waldo 1975; Klein, Petersilia, and Turner 1988, 1990; Lotz and Hewitt 1977; Moore and Miethe 1986), while a smaller proportion of studies find that whites are punished more harshly than minorities (Bernstein, Kelly, and Doyle 1977; Feimer, Pommersheim, and Wise 1990; Myers and Talarico 1986; Peterson and Hagan 1984).

This chapter reviews the existing research regarding the influence of race/ethnicity on sentencing outcomes. Rather than review the results from each individual empirical study (which is the overarching goal of this research), we focus on summarizing and critiquing existing reviews of this voluminous body of literature.

A Synopsis of Prior Reviews of Race-Sentencing Studies

Several notable and frequently cited reviews of the research on race and sentencing have been conducted that summarize what is known about the relationship between race and sentencing outcomes. In this section, we examine five frequently cited

reviews of this literature (Chiricos and Crawford 1995; Hagan 1974; Hagan and Bumiller 1983; Zatz 1987) and one recent "comprehensive" review of the recent research (Spohn 2000a). The purpose of examining these previous reviews is threefold: 1) to gain a sense of the accumulated knowledge regarding the impact of race and ethnicity on sentencing outcomes; 2) to provide a point of reference and comparison for the current research; and, 3) to discuss shortcomings of these earlier reviews that we intend to address.[1] While we focus on these five reviews our comments are not meant to be critical of these authors' work, *per se*; rather the issues we raise are germane to all similar narrative reviews of this research.

In 1974, a very influential review written by John Hagan was published. Hagan reviewed twenty studies *published* between 1928 and 1973 that empirically examined the relationship between "extra-legal variables" (i.e., race, socioeconomic status, sex, and age) and sentencing decisions. Seven of these studies considered the influence of race on sentencing decisions in non-capital cases (excluding one study which used likelihood of conviction as the outcome). For each of these studies, Hagan calculated a measure of association (Goodman and Kruskal's tau-b) between race and sentence severity and a test of statistical significance for each relationship.

Hagan found that the majority of the studies included in his review were methodologically flawed, as most of these studies failed to include controls for legally legitimized factors such as offense type/severity and defendant prior record. Also, few of these studies calculated measures of association; rather, these studies simply relied on tests of statistical significance.

[1] Some of these reviews included studies assessing the influence of race/ethnicity in death penalty decisions; however, we omit a discussion of this research since it is outside of the purview of the current research.

In his re-analysis of these studies, Hagan found that race had a statistically significant relationship with sentence severity in six of the seven studies (with African-Americans being sentenced more harshly in each instance); however, these associations between race and sentence were generally small. After controlling for type of offense, the largest association between race and severity of sentence was 0.08 and the median association was just 0.014. Further, three of the reviewed studies simultaneously controlled for prior record and type of offense; in all three of these studies, there was no relationship between race and sentence, when only offenders without a prior record were examined. In contrast, for offenders with a prior criminal record the relationship between race and sentence severity was "modest" and statistically significant in two of the three studies (see p. 378).

Hagan updated his review in 1983 (Hagan and Bumiller 1983) using a similar method. In this more recent review, Hagan and Bumiller summarized the results of 51 studies, and where possible they calculated measures of association between race and sentence severity. The reviewed studies utilized data as far back as 1864 and as recent as 1977. Hagan and Bumiller organized their review by grouping studies according to whether each study's analysis: 1) used sentencing data on cases adjudicated before or after 1969, and; 2) included controls for offense severity or type and defendant prior criminal record. Hagan and Bumiller's analysis indicates that studies conducted with data before 1968 were only slightly more likely to find indications of racial discrimination than those studies collected with more recent data (56% vs. 54%). Interestingly, later studies that included controls for offense severity and criminal history were more likely than earlier studies employing the same controls to find signs of racial

discrimination (50% vs. 27%). Hagan and Bumiller conclude that overall the relationship between race and sentence is "generally weak" (p. 32-33).

Perhaps the most well known synthesis of the race and sentence literature is Kleck's 1981 review. In this synopsis of the research, Kleck appraised 40 independent studies of non-capital sentencing decisions *published* between 1935 and 1979. Kleck classified each study into one of three categories depending upon what proportion of their findings favored the discrimination hypothesis. Specifically, studies "were characterized as mixed if from one-third to one-half (inclusive) of the findings favored the discrimination hypothesis and as favorable to the hypothesis if more than one-half of the findings favored it" (p. 789). Kleck found that:

> [O]nly eight [of the 40 studies] consistently support the racial discrimination hypothesis, while 12 are mixed and the remaining 20 produced evidence consistently contrary to the hypothesis [of racial discrimination]. . . . However, the evidence for the hypothesis is even weaker than these numbers suggest, since of the minority of studies which produced findings apparently in support of the hypothesis, most either failed completely to control for prior criminal record of the defendant, or did so using the crudest possible measure of prior record–a simple dichotomy distinguishing defendants with some record from those without one. . . It appears to be the case that the more adequate the control for prior record, the less likely it is that a study will produce findings supporting a discrimination hypothesis (p. 789-92).

Based on this analysis, Kleck concludes that "the evidence is largely contrary to a hypothesis of general or widespread overt discrimination against black defendants, although there is evidence of discrimination for a minority of specific jurisdictions, judges, crime types, etc." (p. 799).

The findings of the above-mentioned reviews cast considerable doubt on the premise that systematic racial bias is a primary source of racial disparities in U.S. prisons. This "no discrimination thesis" was also supported by the National Research Council's

Panel on Sentencing Reform, which concluded “the available research suggests that factors other than racial discrimination in the sentencing process account for most of the disproportionate representation of black males in U.S. prisons” (Blumstein et al. 1983:92). Wilbanks (1987) sums up this line of thought most unequivocally by asserting that the idea that the criminal justice is systematically biased against African-Americans is a “myth."

Zatz (1987) presents a different interpretation of these findings in her review of the literature. Unlike most previous reviews, Zatz explicitly focuses on the historical context of each study. Zatz states that there have been four waves of research on sentencing disparities. Wave I was comprised of studies conducted between 1930s and the mid-1960s. In this period numerous studies demonstrated that there was clear and consistent evidence of racial bias against non-whites in sentencing. Research in Wave II reanalyzed earlier studies using more advanced analytic techniques than were available to the original authors. Typically these re-analyses were “interpreted to mean that discrimination was no longer an issue” (p. 73), with the possible exception of the implementation of death sentences in the South. However, these re-analyses also indicated that race may have an indirect discriminatory effect operating through other variables or race interacted with other factors to influence decision making. The third wave of sentencing research was published in the late-1970s and 1980s, using data from the late-1960s and 1970s. Research in this wave indicated that racial discrimination occurred in both overt and more subtle forms in at least some social contexts. The fourth wave is ongoing. Research in this period typically has analyzed data from jurisdictions

with sentencing guidelines. These studies continue to show subtle racial bias against minority defendants.

Based on this interpretation of the literature, Zatz in contrast to earlier reviews concludes that race is a significant determinant of sentencing outcomes. Zatz states that while "it would be misleading to suggest that race/ethnicity is *the* major determinant of sanctioning. . . , race/ethnicity is *a* determinant of sanctioning, and a potent one at that" (p. 87, author's emphasis). Further, Zatz contends with the advent of sentencing guidelines and other structured sentencing mechanisms, racial bias has simply changed form: "Discrimination has not gone away. It has simply changed its form to become more acceptable. . . . [sentencing reform] has caused discrimination to undergo cosmetic surgery, with its new face deemed more appealing" (p. 87). That is, discrimination still exists, but the ways in which discrimination manifests itself has changed. The influence of race now operates through variables considered to be legitimate, or race influences sentencing decisions by interacting with legitimate variables. Thus, Zatz argues discrimination has changed from being overt and direct to being indirect and interactional.

Later reviews continue to find even more compelling evidence of unwarranted racial disparities in sentencing decisions. In a relatively recent review, Chiricos and Crawford reviewed the results of 38 studies *published* between 1975 and 1991. Unlike previous reviews, Chiricos and Crawford devoted a considerable amount of their attention toward explaining why results of sentencing studies vary, in addition to determining whether there are substantial racial disparities. The authors' primary explanatory variables were type of sentencing outcome (the decision to incarcerate versus

sentence length) and structural characteristics of the sentencing jurisdiction (e.g., percent black, unemployment rate, region of sentencing jurisdiction). The authors hypothesized that the impact of race is strongest in Southern jurisdictions, jurisdictions with high percentages of African-Americans in the population and in places with higher rates of unemployment. Because the authors disaggregated research findings by region and structural characteristics, the authors exclude studies focusing on sentencing decisions in the Federal courts.

The 38 reviewed studies produced 145 estimates of the race/punishment severity relationship (only studies examining African-American and white sentencing outcomes were analyzed). Sixty-eight percent of these relationships indicated that African-Americans were sentenced more harshly than whites, and in one-third of the total number of relationships African-Americans were sentenced statistically more harshly than whites. When the authors disaggregated these results by type of sentencing outcome, they found that 85% of the estimates of the relationship between race and imprisonment decisions indicated that African-Americans were punished more harshly than whites, and 52% of these 145 estimates were statistically significant. Whereas 54% of sentence length decisions indicated African-Americans were sentenced more harshly than whites and 20% of these relationships were statistically significant.

Among those estimates that controlled for offense type and prior criminal record, 80% of incarceration decisions and 53% of sentence length decisions found that African-Americans were sentenced more harshly, and 41% and 15% of these estimates were statistically significant respectively. The authors concluded that the evidence “suggests that even when prior record and crime seriousness are controlled for, race is a consistent

and frequently significant disadvantage for blacks when in/out [imprisonment] decisions are considered. At the same time, it appears that race is much less of a disadvantage when it comes to sentence length" (p. 297).

When they disaggregated the relationships by structural features of the sentencing jurisdiction, they found several theoretically relevant associations. In particular, the authors found that race of the defendant impacts imprisonment decisions more consistently in the South than in other regions. The authors' analysis revealed that in the South, African-Americans were more likely to be imprisoned than whites in 88% of the relationships and statistically so in 53% of the relationships; whereas, in non-southern jurisdictions these figures respectively were 76% and 34%. Similarly, Chiricos and Crawford found black defendants were especially more likely to be imprisoned post-conviction than whites in places where blacks comprise a larger percentage of the population and where unemployment was high.

In a very recent review, Spohn appraised all *published* studies concerning the relationship between race/ethnicity and sentencing outcomes that met the following criteria: 1) utilized data on sentences imposed for non-capital offenses during the 1980s and 1990s; 2) reported a measure of association between race/ethnicity and sentence severity; 3) used appropriate statistical techniques; and, 4) included controls for crime seriousness and prior criminal record (p. 453). Based on these criteria, Spohn obtained a sample consisting of 40 studies, 32 of which examined state/local sentencing practices and the eight remaining studies examined Federal sentencing practices. In a manner similar to Chiricos and Crawford, Spohn disaggregated the findings of these studies by

type of sentencing decision (imprisonment decisions, sentence length, sentencing departures) and type of court (State or Federal).

Spohn's review goes beyond earlier summaries of the literature in two significant ways. First, Spohn reviewed sentencing outcomes of Hispanics in comparison to non-Hispanic whites, in addition to comparing sentencing outcomes of African-Americans to whites. Second, she also reviewed findings concerning the indirect and interaction effects of race/ethnicity on sentencing outcomes. Earlier reviews of the literature often made reference to more subtle, indirect and interactive race effects (e.g., see Hagan 1974; Zatz 1987), but heretofore no review had systematically examined the evidence regarding these subtle effects.

In concordance to Chiricos and Crawford's earlier review, when Spohn examined studies using data collected from state and local courts, she found more consistent evidence of unwarranted racial and ethnic disparity in imprisonment decisions than in decisions pertaining to sentence length. This general finding also held for Hispanics in studies examining Federal sentencing decisions. However, at the Federal level, for African-Americans the most consistent evidence of unwarranted sentencing decisions was found in regards to sentence length decisions. Furthermore, while only a small portion of the studies reviewed included analyses of departure decisions, Spohn also found strong evidence of unwarranted disparity in this type of decision in both State and Federal courts.

Spohn's review of the indirect and interaction effects of race/ethnicity uncovered four themes: First, minorities are particularly likely to be treated more harshly when minority status is combined with being male, young, and low socioeconomic status (i.e.,

low income, unemployed, or less educated). Second, process-related factors such as retaining a private attorney or pleading guilty condition the effect of race/ethnicity. Third, at least in sexual assault cases race of the defendant appears to interact with the race of the victim, producing the most severe punishments for African-American defendants who assault whites. Fourth, the influence of race/ethnicity is particularly strong in cases involving drug and less serious offenses.

Spohn concludes:

> [T]he disproportionate number of racial minorities confined in our Nation's jails and prisons cannot be attributed solely to racially neutral efforts to control crime and protect society. . . . Black and Hispanic offenders—and particularly those who are young, male, or unemployed—are more likely than their counterparts to be sentenced to prison . . . Other categories of racial minorities—those convicted of drug offenses, those who victimize whites, those who accumulate more serious prior criminal records, or those who refuse to plead guilty or are unable to secure pretrial release—also may be singled out for punitive treatment (p. 481).

Summary of Findings from Existing Syntheses

Several patterns are noticeable by tracing the evolution of this body of research through these reviews. First, the empirical research in this area has become considerably more methodologically rigorous since Hagan's initial review. It is common for contemporary studies to control for offense seriousness and offender prior record, as well as demographic and oftentimes socioeconomic characteristics of the defendant. While only three studies in Hagan's 1974 review controlled simultaneously for these factors, Spohn's review turned up 40 contemporary studies that at a minimum controlled for offense seriousness and defendant prior record. Furthermore, a sizeable portion of contemporary studies control for potential selection bias issues and examine interactive

and indirect race effects. Such features were for the most part noticeably absence in early studies.

Second, there is increasing evidence that race and ethnicity do influence sentencing decisions. Early reviews found little evidence of racial bias. Typically, the studies included in these early reviews analyzed sentencing length outcomes utilizing a limited number of controls and weak analytic strategies. Early researchers found the primary determinants of sentence severity were offense seriousness and defendant prior record. After accounting for these factors, the overall association between race and sentence was found to be weak and inconsistent. Such findings apparently led many to accept the "no discrimination thesis." Recent studies continue to find that the primary determinants for sanction severity are seriousness of the offense and prior criminal record. However, in contrast to earlier studies, contemporary studies tend to find notable and fairly consistent indications of unwarranted racial/ethnic disparity in various aspects of the sentencing process, especially in regards to imprisonment decisions. Moreover, although relatively few studies have examined discretionary outcomes (e.g., downward departures from guidelines), those studies which do find consistent evidence of unwarranted racial/ethnic disparity.

The increased consistency of findings of unwarranted racial/ethnic disparities in recent years appears anomalous given the apparent racial progress America has experienced in the past several decades. How can this trend of increasingly consistent evidence of unwarranted racial disparities be explained? It seems likely that the best explanation for this anomaly is that researchers have focused their collective attention to those sentencing contexts most likely to find such disparities. Hagan and Bumiller

(1983) reach a similar conclusion, when they note that recent researchers in this area: "have focused more selectively on those structural and contextual conditions that are most likely to result in racial discrimination" (p. 21).

Perhaps the best evidence of this selectivity is found by noting the shift in the types of sentencing outcomes analyzed. For example, in Hagan's original review the overwhelming majority of the relationships between race and sentence used sentence length as the outcome variable and a much smaller proportion used sentence type (e.g., prison or probation) as the outcome. Over time the distribution in the types of sentencing outcomes analyzed has changed. In Hagan and Bumiller's review the distribution in types of sentencing outcomes is appropriately a 60/40 split with sentence length outcomes still being the majority; however, in Spohn's review sentence length outcomes comprised a minority of the relationships. This shift is significant because the above reviews indicate that there is considerably less evidence of unwarranted disparity in regards to sentence length outcomes than other types of sentencing outcomes. Thus our gain in knowledge regarding the influence of race/ethnicity on sentencing outcomes appears to have affected the manner in which research in this area is being conducted. It appears that these changes are at least partially responsible for the increasingly consistent finding of unwarranted racial disparity.

Critique of Existing Reviews

While the above-mentioned reviews have significantly advanced our understanding of the relationship between race/ethnicity and sentencing decisions, we believe that these summaries have several limitations that need to be addressed (see Table

1). There are five primary weaknesses exhibited in previous reviews: 1) a focus on statistical significance of individual outcomes, instead of the magnitude and direction of observed effects; 2) non-comprehensive search strategies; 3) (implicit) equal weighting of studies regardless of sample size or methodological rigor; 4) the inclusion of studies based on the same or overlapping data sets; and, 5) a focus only on African-American sentencing outcomes in comparison to whites (i.e., omit ethnicity). By pointing out these shortcomings in prior syntheses, we do not intend to belittle the work of other researchers; rather, we believe the field will not advance unless we engage in candid discussions of such shortcomings and begin to address these issues. In this spirit, we critique several of the studies reviewed above.

Table 1. Shortcomings of Previous Reviews

Review	Statistical Significance[a]	Equal Weight[b]	Published Studies[c]	Dependent Studies[d]	Race Only[e]
Hagan (1974)	No	Yes	Yes	No	Yes
Hagan & Bumiller (1983)	No	Yes	Yes	No	Yes
Kleck (1981)	Yes	Yes	Yes	No	Yes
Chiricos and Crawford (1995)	Yes	Yes	Yes	Yes	Yes
Spohn (2000a)	Yes	Yes	Yes	Yes	No

[a] Review focused primarily on the statistical significance of individual results, instead of the magnitude and direction of the influence of race/ethnicity .
[b] Review implicitly gives equal weight to included studies.
[c] Review focused on published studies.
[d] Review includes multiple studies based on the same or overlapping data sets.
[e] Review focused only on the influence of race (i.e., African-American vs. white).

With the notable exceptions of the reviews led by Hagan (Hagan 1974; Hagan and Bumiller 1983), existing reviews focus predominantly on the statistical significance of race/ethnicity, instead of the more meaningful criteria of magnitude and direction of racial/ethnic differences. However, reviews focusing on statistical significance assume,

at least implicitly, that failure to reject the null hypothesis supports a null conclusion. However, failing to find a statistically significant relationship between two variables does not mean that no relationship exists—it simply means that the evidence is not strong enough to reject the null hypothesis of no relationship.

As a result, instead of addressing the issue of fundamental importance: What is the magnitude and direction of unwarranted disparities between minorities and whites? Research syntheses focusing on statistical significance address fundamentally flawed questions, such as: What proportion of studies find statistically significant race/ethnicity effects? This focus inhibits these reviews from being able to determine whether minorities are 50%, 20%, or 1% more likely to be imprisoned than similarly situated whites–all of which could be statistically significant given the large sample sizes used in many studies, especially recent studies.

Moreover, by simply finding what proportion of studies reveal statistically significant race/ethnic differences, implicitly these studies give equal weight to each study regardless of the study's sample size, precision of estimation, or methodological rigor. It would seem more reasonable to give more weight to studies with more precise estimates of the relationship between race/ethnic and sentence severity (typically studies with larger samples). It also seems reasonable to take into account the methodologically rigor of each study. That is, a prudent question to address is whether the magnitude of unwarranted disparity is systematically smaller or larger in studies with more methodological rigor than less rigorous studies.

All of the extant reviews are non-comprehensive as they focus on published studies—opening the door for publication bias to creep into these reviews.[2] Researchers in other areas of study have repeatedly demonstrated that published studies are a biased subset of an area of research, as published studies are more likely to find statistically significant results than unpublished research (Greenwald 1975; Lipsey and Wilson 1993; Smith 1980). Thus, existing reviews run a substantial risk of biasing their results by focusing only on published research. A more judicious approach would be to include all available research meeting explicit eligibility criteria, regardless of publication status.

Another questionable practice evident in several extant reviews is the inclusion of multiple studies based on the same or overlapping data sets and hence, in essence, double-count the same study. Specifically, the reviews of Chiricos and Crawford and Spohn include several of such overlapping studies. For instance, Spohn's review includes two studies investigating the effect of race/ethnicity on sentencing outcomes in the Federal system that use virtually identical data sets (Langan 1999; United States Sentencing Commission 1995 both of which analyze 1994 data on 14,000 Federal drug trafficking cases). Similarly, Spohn's review includes a series of studies conducted by Myers analyzing Georgia sentencing data in the late 1970s and early 1980s (Myers 1987, 1989; Myers and Talarico 1986). Spohn simply notes that because her review is intended to be comprehensive and because none of these studies used *exactly* the same the data, the inclusion of these studies is justified (see p. 454-55). The problem with this approach is that these samples are not independent, even if they are not exactly the same. In other words, the substantial overlap between these data sets makes further analysis of these

[2] Some of these reviews (Hagan and Bumiller 1983; Spohn 2000a) do include a few unpublished studies; however, given the researchers' search strategy clearly the focus was on locating published works.

data redundant; knowing the results of the first study enables one to predict the results of later studies using the same data. In essence, Spohn double-counts and sometimes triple-counts the same study (or data set).

Furthermore, "the mental algebra" required in narrative syntheses can also lead to findings inconsistent with the research. For example, in Kleck's (1981) review of the research he makes several decisions that appear arbitrary and susceptible to bias. A case in point is Kleck's decision to classify studies as supporting the discrimination hypothesis based upon the proportion of results finding race differences. Specifically, Kleck characterizes studies "as mixed if from one-third to one-half (inclusive) of the findings favored the discrimination hypothesis and as favorable to the hypothesis if more than one-half of the findings favored it" (p. 789). Clearly, this categorization is arbitrary. Further, Kleck never defines how "favorable to discrimination hypothesis" is operationalized.

Some might not be convinced that these shortcomings have meaningful repercussions for these reviews. However, the work of Mann (1994) and Cooper and Rosenthal (1980) discussed in the previous chapter demonstrate that such weaknesses run a substantial risk of leading narrative reviews to reach conclusions inconsistent with the empirical research. Given the shortcomings of existing syntheses, we believe that there are reasonable grounds to question whether existing reviews have accurately captured the cumulative findings of the empirical research in this area.

CHAPTER 3. RESEARCH DESIGN AND ANALYTIC STRATEGY

Research Methodology

Glass (1976) categorizes research into three types: primary, secondary, and meta-analysis. Primary research concerns the analysis of original data. Secondary research is the re-analysis of data for the purpose of answering the original research question with more advanced analytic techniques or addressing new questions with previously analyzed data. Glass refers to meta-analysis as "the analysis of analyses . . . [T]he statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings" (p. 3). Glass argues that meta-analysis provides a method for summarizing the results from a large body of research in a manner that permits *knowledge* to be extracted from a mass of *information* provided by individual studies.

The key to summarizing the results from a quantitative body of research is to standardize results from each study in a manner that facilitates comparisons across studies. Meta-analysis accomplishes this important task by utilizing "effect sizes." While there are many different types of effect sizes, the common goal of these various measures is to create a quantitative scale capable of capturing variation in the direction, magnitude, or both of results from a body of research (Lipsey and Wilson 2001). These effect sizes are then utilized in data analyses, in much the same way as other dependent and independent variables.

Lipsey and Wilson explain the method of meta-analysis by comparing it to survey research:

> Meta-analysis can be understood as a form of survey research in which research reports, rather than people, are surveyed. A coding form (survey protocol) is

> developed, a sample or population of research reports is gathered, and each research study is 'interviewed' by a coder who reads it carefully and codes the appropriate information about its characteristics and quantitative findings. The resulting data are then analyzed using special adaptations of conventional statistical techniques to investigate and describe the pattern of findings in the selected set of studies (p. 1-2).

Thus, one way to think of meta-analysis is as a survey of the existing research, where typically each study's empirical results are treated as scores on the dependent variable(s) and features of the study are coded as independent (or moderator) variables. Once a database containing scores on the dependent and independent variables has been created, these data are analyzed in a manner similar to conventional analyses.

According to Cooper and Hedges (1994), there are five major steps in conducting a meta-analysis: 1) formulating the research question(s); 2) searching the literature; 3) coding empirical studies; 4) analysis and interpretation; and, 5) public presentation. This chapter addresses the first four of these steps. The resulting meta-analysis is designed to remedy the shortcomings of previous reviews of this research and to be a truly comprehensive synthesis of race/ethnicity and sentencing research.

Research Questions and Hypotheses

The specific aim of this study is to conduct a systematic and comprehensive meta-analysis that examines: 1) whether race and ethnicity are related to sentencing outcomes, independent of offense seriousness and defendant criminal history; 2) whether the results from existing empirical studies are systematically related to methodological and sample variations; and, 3) in which contexts are unwarranted racial/ethnic disparities most likely to occur (e.g., jurisdictions without structured sentencing, Southern jurisdictions). These

research questions combined with the findings of earlier are used to generate the following hypotheses:

H_1: *Generally*, racial and ethnic minorities are sentenced more harshly than whites, independent of offense seriousness and criminal history.

H_2: Unwarranted racial/ethnic disparity is smallest in analyses that control for factors related to both race/ethnicity and sentence severity, and utilize precise measures of these variables. That is, analyses that employ interval-level or multiple dichotomous measures of criminal history, instead of a single dichotomy, and analyses that measure offense severity utilizing ordinal-level offense severity ratings rather than common law measures of offense type (e.g., drug, violent, property offenses) produce smaller estimates of unwarranted racial/ethnic disparities.

H_3: Part of the variability in research findings is attributable to differences in sample characteristics.

H_4: Unwarranted racial disparity is largest in those contexts where sentencing discretion is greatest (i.e., jurisdictions without sentencing guidelines).

Eligibility Criteria for Meta-Analysis

The eligibility criteria for inclusion in the present research were that each study had to: 1) be conducted using cases sentenced in the United States; 2) examine sentencing outcomes in criminal courts (i.e., juvenile court adjudication/sentencing outcomes are excluded); 3) incorporate simultaneous controls for both offense seriousness and criminal history; 4) measure the direct influence of race/ethnicity on sentencing outcomes; 5)

examine sentencing outcomes unrelated to death penalty decisions; and, 6) all research must be made available through yearend 2002. Note that eligible studies may be published or unpublished. Moreover, studies need not be primarily concerned with the influence of race/ethnicity on sentencing outcomes; regardless of the study's focus, as long as the analysis measured the direct influence of race/ethnicity on some sentencing decision, the study was eligible for inclusion in this synthesis. It is also important to note studies that examined *only* the indirect or interactional effect of race/ethnicity on sentence severity were not included in this review.

Studies examining the influence of race/ethnicity on sentencing outcomes in capital and juvenile (i.e., non-criminal) offenses are excluded, as we believe the inclusion of such studies would lead to comparisons of apples and oranges. That is, mixing sentencing outcomes regarding capital offenses with more mundane offenses could obscure salient differences between the manner in which race influences sentencing outcomes between capital offenses and non-capital offenses. Similarly, we exclude studies that examine sentencing outcomes of in juvenile courts, as juvenile courts' *parens patriae* philosophical orientation introduces a host of issues (e.g., needs of the child) usually not considered in criminal courts. We believe that separate meta-analyses of these types of court decisions are more appropriate than mixing these various types of court cases.

Moreover, only those studies that control for offense seriousness (or offense type) and criminal history have been included. This criterion is necessary for several reasons. First and foremost, the central point of disagreement between disparity and discrimination explanations of the over-representation of racial and ethnic minorities in

the U.S. criminal justice system concerns the effect of race after these factors have been taken into account. Studies that do not take these factors into account do not shed any light on the veracity of these competing explanations. That is, studies that do not at a minimum control for seriousness of current offense and extent of prior defendant criminal history, do not eliminate the two primary factors believed to account for observed differences between minorities and whites on sentencing outcomes. Second, a considerable body of literature indicates that these two variables are consistently the most important determinants of sentencing outcomes (Bernstein et al. 1977; Blumstein et al. 1983; Chiricos and Waldo 1975; Kleck 1981; Kramer and Ulmer 1996; Souryal and Wellford 1999; Wooldredge 1998). The existing literature also has shown that there are meaningful differences between racial/ethnic groups on these two factors; for example, African-American defendants tend to have longer criminal histories than whites (Albonetti 1997; Miethe and Moore 1986; Petersilia and Turner 1987). Thus, any study that does not, at a minimum, include controls for these factors runs a high risk of introducing specification error into its estimate of the relationship between race/ethnicity and sentencing outcomes, and therefore has been excluded.

This meta-analysis focuses on five types of sentencing outcomes: 1) imprisonment decisions, 2) length of incarcerative sentence, 3) simultaneous examinations of imprisonment and sentence length decisions, 4) discretionary lenience, and 5) discretionary punitiveness. Imprisonment and length of incarcerative sentence decisions are self-explanatory. The third type of sentencing outcome, what we refer to as simultaneous examinations of imprisonment and sentence decisions, typically analyze sentencing decisions using ordinal scales with probation sentences being the least severe

sentence, short-term incarcerative sentences being the next level of severity, and longer-term incarcerative sentences being progressively more severe on the ordinal scale. For example, Harig (1990) analyzes an 11-point sentence severity scale "that considers an incarcerative prison term (11) the most severe possible sentencing outcome followed, in decreasing severity, by (10) jail, (9) time served, (8) jail and probation, (7) probation and fine, (6) probation, (5) fine with conditions, (4) fine, and (3) conditional discharge, (2) unconditional discharge, and (1) other lesser sentences" (p. 106). Discretionary lenience refers to sentencing outcomes where a sentencing authority, typically the judge, sentences a defendant to a punishment more lenient in some manner than ordinary; e.g., downward departures from guidelines, stays of sentence, and so forth. Conversely, discretionary punitiveness refers to sentencing outcomes where a sentencing authority sentences a defendant to a sanction more harsh than ordinary; e.g., upward departures, enhanced sentencing provisions for eligible repeat offenders, consecutive sentences (instead of concurrent sentences), and so forth. By labeling these outcomes "discretionary lenience" and "discretionary harshness," we do not mean to imply that these cases were inappropriately lenient or punitive; rather, these terms are used to denote sentencing outcomes involving punishments that differ from the standard sentence in some meaningful way.

The decision to include both published and unpublished studies invariably leads to a concern of how rigorous studies must be in order to be eligible for inclusion. The concern is that unpublished studies may be of lower methodological rigor than published studies. The majority of unpublished studies in this body of research, however, were doctoral dissertations, which generally displayed a high level of methodological rigor.

Moreover, there is a tremendous amount of variation in the rigor of published studies (Wooldredge 1998). We believe the requirement that all studies include statistical controls for offense seriousness and defendant criminal history provides a reasonable lower limit of methodological rigor. Additionally, key features of each study's methodology, and data analysis have been coded, such as the number and types of other control variables included in the analysis (e.g., study includes controls for defendant socio-economic status, type of attorney, method of disposition). This information is then utilized to test whether methodological rigor is systematically associated with the size of racial/ethnic differences in sentencing outcomes.

Another key decision is deciding which study should be included when several studies analyze the same data. One of our primary criticisms of existing reviews is that they include several studies based on the same or very similar data sets, which leads to double-counting of what is essentially the same study. When multiple studies are encountered that rely on the same data, the decision of which study should be included in this meta-analysis was based on the following criteria (listed in level of priority):

1) Codeability—Invariably some analyses are uncodeable, typically, because they lack important information (e.g., standard deviations or sample size are not reported) or the type of analysis does not lend itself to effect size coding (e.g., structural equation models). Thus, the first criterion is that studies reporting results in a manner unsuitable for effect size coding are excluded from the analysis.

2) Context specificity—Because a focal point of this research is to assess the degree to which unwarranted racial/ethnic disparity varies systematically with features of the sentencing context, studies that analyze data in the most context specific manner are

given highest priority. For example, if two studies analyze the same data set, one of the studies disaggregates its analyses by contextual features such as time period or place, while the other study simply pools all of the data together without regard to these contextual features, the first study was selected for inclusion in this meta-analysis as it is more context specific.

3) Methodological rigor—Those studies that use the most appropriate data analytic technique and included a greater number of control variables were preferred over other studies utilizing the same data.

4) Sample size—Studies with larger sample sizes were given priority.

Search Strategy

It is well documented in the meta-analytic literature that relying solely on published studies may produce a biased sample and is therefore inadequate (Callaham 1998; Greenwald 1975; Smith 1980); hence, unpublished studies as well as published studies were included in this analysis. Our search strategy included examination of: 1) bibliographies from existing syntheses; 2) references contained in eligible studies; 3) computerized bibliographic databases (e.g., NCJRS, Criminal Justice Abstracts, Sociological Abstracts, Social Science Citation Index, PsycINFO, ERIC); 4) hand searches of select relevant journals (Criminology, Journal of Quantitative Criminology, Justice Quarterly); 5) dissertations via Dissertation Abstracts; and, 6) conference programs through online searches of Conference Papers Index and hand searches of relevant conference proceedings (e.g., American Society of Criminology).

In order to obtain an expansive search of computerized databases, we utilized a number of keywords and numerous combinations of keywords. Our search of computerized databases used the following keywords in multiple combinations: sentencing, judicial sentencing, judicial discretion, sentencing discretion, sentencing reform, sentencing research, court research, sentencing disparity, sentencing discrimination, race, ethnicity, African-American(s), Black(s), Hispanic(s), Latino(s), Native American(s), Asian(s), racial discrimination, racial bias, and racial disparity. By using such broad keywords we believe that a negligible number of studies were missed.

In addition, we contacted each state's sentencing body to determine whether these organizations have internally evaluated their jurisdiction's sentencing practices in regards to unwarranted racial/ethnic sentencing disparity. Specifically, we utilized contact information listed in the National Association of State Sentencing Commissions' newsletter to establish a mailing list. We then contacted each of these identified sentencing bodies to inquire about research conducted in their jurisdiction regarding racial/ethnic disparity in sentencing outcomes.

Once a prospective study was identified, a *preliminary* screening was made on the basis of title, abstract, and any other available information. We attempted to retrieve a full copy of all studies that were not clearly disqualified based on this preliminary review. That is, the preliminary review of each study's title and abstract was used only to disqualify studies clearly failing to meet the established eligibility criteria. The full versions of these studies were then reviewed to determine final eligibility.

Effect Size Coding

An effect size was calculated for each minority-white contrast from all eligible, independent sentencing contexts. These effect sizes are the dependent variable for this meta-analysis. Effect sizes were coded in manner such that *positive effect sizes indicate minorities were punished more harshly than whites*, and negative effect sizes indicate minorities were punished less harshly than whites. Specifically, the (logged) odds-ratio was chosen as the preferred effect size for outcomes analyzing dichotomous dependent variables (e.g., the likelihood of receiving a sentence involving incarceration), whereas the standardized mean difference effect size was chosen for outcomes with interval-level or continuous dependent variables (e.g., length of sentence). The odds-ratio effect size (ES_{or}) is defined as:

$$ES_{or} = \frac{P_m / 1 - P_m}{P_w / 1 - P_w} \tag{1}$$

where p_m is the probability of the event (e.g., imprisonment) for minorities and p_w is the probability of the same event for whites. The standardized mean difference effect size (ES_d) is defined as:

$$ES_d = \frac{\bar{X}_m - \bar{X}_w}{S_{pooled}} \tag{2}$$

where $\bar{X}_m$ is the minority group mean, $\bar{X}_w$ is the white group mean, and s_{pooled} is the pooled within groups standard deviation, defined as:

$$S_{pooled} = \sqrt{\frac{(n_w - 1)s_w^2 + (n_m - 1)s_m^2}{(n_w - 1) + (n_m - 1)}} \tag{3}$$

where s_w^2 is white group variance, s_m^2 is the minority group variance, n_w is the white group sample size, and n_m is the minority group sample size.

Because this meta-analysis is concerned with the influence of race and ethnicity, after prior criminal record and current offense seriousness have been taken into account, the actual calculation of the effects sizes had to be modified slightly. Most of the effect size estimates were derived directly from the primary author's multivariate analyses;[3] for example, effect sizes for dichotomous outcomes were most often taken straight from the primary author's multivariate logit analyses (logistic regressions) when available, as the results from these analyses are already reported as (logged) odds ratio effect sizes. Effect sizes for continuous outcomes also were derived from primary author's multivariate ordinary least squares analyses. For this type of outcome, the numerator in equation 2 was replaced by the unstandardized race/ethnicity regression coefficient, as this regression coefficient reflects the difference between minorities and whites after other factors have been taken into account, which is equivalent to the numerator in equation 2.

In more than a few instances, the primary authors analyzed a dichotomous outcome with ordinary least squares (OLS) regression. While this practice was common prior to the proliferation of computer programs capable of performing analyses of limited dependent variables, currently this practice is considered unacceptable, because this practice violates several assumptions of OLS regression. One of the most serious problems with this practice is that the resulting regression coefficients are inefficient; however, the regression coefficients generally are not biased (Aldrich and Nelson 1984; Allison 1999; Long, 1997). That is, the regression coefficients still are unbiased

[3] Some multivariate analyses were performed on correlation matrices or contingency tables provided by the primary authors.

estimators of the average difference in probability of receiving some sentencing outcome by race/ethnicity, but the standard errors associated with these regression coefficients are biased due to heteroskedascity.

Because our calculations of effect size do not depend on the primary authors standard errors and because the regression coefficients are unbiased estimators of the race/ethnicity's influence on sentence severity, we decided to code the results from these analyses rather than discard these studies' results. For these studies, we calculated the standardized mean difference effect size as described above, and then transformed this effect size onto the odds ratio scale by multiplying this effect size by $\frac{\pi}{\sqrt{3}}$ (for a discussion of this conversion see Hasselblad and Hedges 1995, or Lipsey and Wilson 2001: 198). We flagged each of these transformed effect sizes in the data set in order to test whether these effect sizes were systematically different than the other effect sizes.[4]

It is important to note that one study may report multiple *independent* minority-white contrasts. For example, Welch, Spohn, and Gruhl (1985) analyzed sentencing data from six *different* jurisdictions and analyzed the data from each jurisdiction separately. This procedure produced six independent estimates of the relationship between race/ethnicity and sentence severity; all six of these independent effect sizes were coded into the current study's data set. From this example it should be clear that *the primary unit of analysis in this research is the independent minority/white contrast, not the study*—as a study may report several independent minority/white contrasts.

[4] It is also worth noting that in a few instances dichotomous outcomes were analyzed by the primary authors using probit regression. The results from these studies were transformed onto the odds ratio scale by multiplying unstandardized regression coefficients by 1.81 (see Long, 1997: 48).

A study may also report multiple *dependent* minority-white contrasts. There are several ways of producing multiple dependent contrasts. For example, a study could use the same sample (or a sub-sample of the total sample) to estimate the influence of race/ethnicity on multiple measures of sentence severity (e.g., imprisonment and sentence length decisions). As another example of a dependent contrast, a study could use the same sample to estimate the influence of race/ethnicity on the same measure of sentence severity (say imprisonment decisions) but across multiple types of offenses (e.g., violent and property offenses). In both of these examples a statistical dependency is caused by using the same sample to produce multiple measures of the influence of race/ethnicity.

A third example of a common type of dependent contrast occurs when a study compares sentencing outcomes of African-Americans to whites and compares sentencing outcomes of Hispanics to the same group of whites. This dependency may be less obvious than the other two examples; in essence, this situation creates a statistical dependency between the African-American/white contrast and the Hispanic/white contrast, because the same comparison group, namely whites, is used in both contrasts.

For the above discussion it should be clear that we have utilized a limited definition of "statistical independence"; in that, we have only attempted to account for statistical dependencies caused by the inclusion of the same case (or person) in multiple contrasts. This is a limited definition of statistical independence because other types of statistical dependencies between contrasts may also exist. For example, perhaps there are statistical dependencies among studies conducted by the same author(s). We have not attempted to take into account other sources of potential statistical dependence, as we

believe that these other potential sources of dependence are small, and therefore will not bias the results of the present research.

Treatment of Dependent Contrasts

An important decision in meta-analysis concerns how the meta-analyst decides to handle dependent effect sizes, as the inclusion of dependent effect sizes into an analysis would violate the statistical independence assumption crucial in many types of data analysis (such as the multiple regression procedure utilized in this analysis). The three types of dependent effect sizes (described above) were handled in the following manner. When studies reported multiple measures of sentence severity, either by disaggregating results by type of sentence outcome (e.g., imprisonment and sentence length decisions) or disaggregating results by type of offense (e.g., person, property, drug offenses), we created independent effect sizes by utilizing two complimentary procedures.

First, we created an *overall* measure of unwarranted disparity by calculating the weighted mean effect size from the multiple dependent effect sizes, weighing by the inverse standard error of each outcome. For example, in contexts that analyzed more than one type of sentencing outcome, most commonly imprisonment decisions and decisions regarding length of incarcerative sentence, then this overall effect size measure is the weighted average of the effect sizes from these separate analyses. When the various effect sizes are measured on different scales (e.g., odds ratio and mean difference), we converted effect sizes onto the odds ratio scale (as this was the most common metric) following the conversion factor provided in Lipsey and Wilson (2001: 198). By contrast, in contexts where only one sentencing outcome was examined and

hence only one effect size was computed, the overall measure of unwarranted disparity is simply the single available effect size.

Analyses reporting multiple offense types, typically, begin by reporting results of a pooled model (e.g., all offense types are included) then in subsequent analyses offense specific results are reported. For example, Souryal and Wellford (1999) analyze the influence of race on imprisonment decisions (among other analyses). These authors first examined the influence of race on imprisonment decisions by pooling all cases regardless of offense type, and then these authors analyze the influence of race on imprisonment decisions disaggregated by offense type (drug, person, and property). The overall measure of unwarranted disparity in such analyses is based on initial pooled (mixed offense type) model. By contrast, in studies not reporting the results of a pooled offense type model (i.e., only offense specific models), then the overall effect size is the weighted average of these separate offense specific effect sizes.

We believe this overall measure of unwarranted disparity is very important, because in most primary research separate sentencing outcomes often are treated as disconnected outcomes, and as a result one fails to gain a sense of the overall influence of race/ethnicity across all sentencing outcomes (and offense types). By combining these separate sentencing outcomes into one measure, we believe the overall influence of race/ethnicity is more accurately portrayed. This overall effect size measures is the primary dependent variable in the analyses that follow.

In order to clarify the calculation of the overall effect size, we will describe the computation of the overall effect size measure for Souryal and Wellford (1999) as an example. In this empirical examination of sentencing in the State of Maryland, these

authors examine unwarranted racial/ethnic disparity in sentencing decisions regarding imprisonment ("in/out") and length of incarcerative sentences. Furthermore, after presenting the results from models pooling all offense types, these authors disaggregated their analyses by type of offense (see Table 2). In all, eight effect sizes were coded from Souryal and Wellford's analyses (see Table 2). The overall effect size in this study was calculated by taking the weighted average of the effect sizes coded from the analysis imprisonment decisions using all offense types and the analysis of sentence length decisions using all offense types; i.e., the effect sizes computed from contrasts 1 and 5 from Table 2 were averaged. An odds ratio effect size was computed from imprisonment decisions, whereas a standardized mean difference effect size was computed from the analysis of the sentence length. Note that it is important to weight the effect sizes, as each effect size is based on a different number of cases; further, because these effect sizes were measured on different scales, we converted the standardized mean difference effect size onto the odds ratio scale before taking the weighted average of the effect sizes. As Table 2 shows the odds ratio effect size calculated from their analysis of imprisonment decisions in this study was based on a sample of 75,929, whereas the standardized mean difference effect size computed from the sentence length analysis was based on 52,627 cases.

Table 2. Effect Sizes Computed from Souryal and Wellford (1999)

Contrast	Contrast Type	N
	Imprisonment Decisions	
1	--All Offense Types	75,959
2	--Violent (Person) Offenses	25,780
3	--Drug Offenses	39,761
4	--Property Offenses	15,418
	Sentence Length Decisions	
5	--All Offense Types	52,627
6	--Violent (Person) Offenses	15,112
7	--Drug Offenses	27,589
8	--Property Offenses	9,926

Second, we conducted separate analyses of each type of sentencing outcome and type of offense. That is, separate effect sizes were computed for imprisonment decisions, sentence length decisions, and so forth; separate effect sizes were also computed for the various types of offenses (property, drug, and violent offenses). These separate effect size analyses are an important augment to the general measure of unwarranted disparity (discussed above) as the overall measure of disparity has the potential to obscure any differential effects of race/ethnicity between sentencing outcomes or offense types.

To illustrate the process of calculating outcome and offense specific effect sizes, we continue to use Souryal and Wellford (1999) as an example. For example, in this study, the imprisonment specific effect size is simply the effect size computed from the first contrast (see Table 2), whereas the sentence length specific outcome is the effect size computed from the fifth contrast. The offense specific effect sizes are computed by taking the weighted average of effect sizes from each type of offense. That is, the drug offense effect sizes in this study were computed by taking the weighted average of the third (imprisonment decisions involving drug offenses) and seventh (sentence length decisions involving drug offenses) contrasts. Once again, because these effect sizes are

on different scales (odds ratio vs. standardized mean difference) before taking the weighted average the standardized mean difference effect size was converted onto the odds ratio scale.

Lastly, when studies reported the results of analyses contrasting multiple minority groups to white, these dependent effect sizes were made statistically independent by conducting all analyses separately for each minority/white contrast. That is, all analyses comparing sentencing outcomes of African-Americans to whites were conducted separately from those contrasting sentencing outcomes of Hispanics to whites. As long as effect sizes from each minority-white contrast are kept separate, this source of dependency is remedied.

Moderator Variable Coding

Each effect size is accompanied by a set of variables that describe its particular characteristics and the context from which it comes. Because a primary emphasis of this research is to explain variability in effect sizes (i.e., study results), our selection of moderator variables was vital. We have attempted to select variables that prior research indicates are important predictors of sentence severity, especially predictors that may be correlated with race/ethnicity.

The existing research indicates that, at the individual-level, offender characteristics have important relationships with sentencing outcomes. Specifically, age and sex of defendant have been found in some research to be important predictors of sentence severity. For instance, while the research is clearly not uniform in this regard, several studies have found women receive more lenient sanctions than men (Bernstein

1979; Daly and Tonry 1997; Frazier and Bock 1982; Steffensmeier 1980; Mustard 2001). Similarly, research by Steffensmeier et al. (1998), Spohn and Holleran (2000), Zatz and Hagan (1985) among many others, have found sentence severity is related to age of defendant. Such findings suggest that these characteristics are potentially important moderator variables; that is, differences between studies on proportion of females in the sample and mean age of sample may contribute to differences in results. Therefore, we coded these factors from each sample in order to capture between study differences on these offender characteristics.

Socio-economic status (SES) is another offender characteristic demonstrated to have a positive relationship to sentence severity. Chiricos and Bales (1991) review the empirical research regarding this association. These authors found that SES, as measured by unemployment status, had a statistically significant relationship to sentence severity in the majority of studies reviewed, even after these studies controlled for other factors. This relationship, however, was stronger in analyses that used imprisonment decisions as a measure of sentence severity. Furthermore, in their analyses Chiricos and Bales (1991:719) found that "unemployment had a significant, substantial, and independent impact on the decision to incarcerate," after taking into account offense severity, prior record, and other factors. Other measures of SES, such as class status measured in relation to means of production (Hagan and Parker 1985) and ordinal measures of SES (i.e., low vs. high) (Jankovic 1978), also have been found to be related to sentence severity. Furthermore, it is well-established that SES is correlated with race/ethnicity with African-Americans and Hispanics *generally* having lower SES levels than whites. This suggests that studies that control for defendant SES may obtain a more accurate

measure of the influence of race/ethnicity. That is, these studies disentangle the influences of SES from those of race/ethnicity; and therefore, the effect sizes from these studies may be systematically different than studies which do not control for SES. Thus, we coded which studies included measures of defendant SES as an independent variable. For our purposes, defendant socioeconomic status could be measured in a number of ways; the most common measures of SES involved defendant employment status, income, or education level.

Methodological concerns have been a continuing issue in this body of research. Many scholars have been highly critical of the methodological rigor exhibited in this research (Kleck 1981; Wilbanks 1987; Wooldredge 1998). Much of this concern has revolved around the inclusion and adequacy of measures of defendant prior record and offense seriousness. This concern is well founded as these two variables consistently have been found to be the most salient determinants of sentence severity. Kleck (1981), for example, complains: "[of] studies which produced findings apparently in support of the [racial discrimination] hypothesis, most either failed completely to control for prior criminal record of the defendant, or did so using the crudest possible measure of prior record—a simple dichotomy distinguishing defendants with some record from those without one" (p. 789). Further, Kleck states that: "It appears to be the case that the more adequate the control for prior record, the less likely it is that a study will produce findings supporting a discrimination hypothesis" (p. 792). Kleck apparently believes that

ordinal or interval measures of prior record will be stronger predictors of sentence than simple dichotomies, and these more adequate controls will reduce the magnitude of unwarranted racial disparities.[5]

Similarly, Kramer and Steffensmeier (1993) contend that common-law measures of offense type (i.e., violent, property, drug) are "too imprecise to provide a meaningful control for offense severity" (p. 358). That is, some researchers apparently believe that within any nominal category of common-law offense types there is too much variation for such measures to serve as rigorous controls. Therefore, more rigorous measures of offense seriousness such as guideline offense score are necessary in lieu of, or in addition to, common-law categories.

The implication of these criticisms is that once more adequate controls for prior record and offense seriousness are utilized, differences in sentence outcomes by race/ethnicity will be attenuated. In order to test this assertion, we created moderator variables designed to reflect the precision of measurement for prior record and offense seriousness. Specifically, studies that used *one* simple dichotomy as a measure of prior record were distinguished from studies that used either multiple dichotomous measures or ordinal/interval level measures of prior criminal record. We employed a broad definition of what constitutes a control measure for prior record, such as prior arrests, convictions, incarcerations and so forth. Likewise, we classified each study's offense seriousness control measure into three categories of increasing precision: 1) studies that control for

[5] Empirical research addressing the relationship between sentencing outcomes and various measures of prior criminal record have not uniformly found that interval level measures of prior record are more strongly related to sentencing outcomes. Nelson (1989) for instance found that "a variety of criminal record scores [including both dichotomous and interval-level measures] was equally effective at predicting incarceration for persons" (p. 350). By contrast, Welch et al. (1984) found that dichotomous measures based on arrests or convictions displayed small correlations with sentence severity.

offense severity using common law offense types (i.e., violent, property, drug, public order offenses) or examine only one common law type of crime, 2) studies that utilize severity of offense ratings (e.g., guideline offense severity scores) or employ offense categories with more precision than common law categories (e.g., analyze only armed robberies), and 3) studies which utilize *both* of the foregoing, which we consider as the highest level of precision.

Additionally, we coded whether the analysis included controls for other measures of offense severity. Specifically, we coded two indicator variables; the first denotes studies that controlled for the presence/use of a weapon, and the second flags studies that controlled for victim injury. Lastly, we coded the total number of variables related to offense seriousness that were entered into each analysis. These variables could include type of offense, ratings of offense seriousness, number of charges/convictions, original charge seriousness (or type), presence of weapons, victim injury, and so forth (see Appendix A for a copy of the coding manual).

Scholars such as Wilbanks (1987) have pointed out other methodological issues that arguably moderate the magnitude of unwarranted racial/ethnicity disparities. In particular, Wilbanks argues forcefully that studies which include controls for factors associated with both sentence severity and race yield small (perhaps trivial) estimates of unwarranted racial disparity. That is, Wilbanks suggests that by omitting variables correlated with race and sentence, researchers have committed a specification error, causing such studies to systematically overestimate the influence of race on sentencing outcomes. Based on this argument, Wilbanks concludes "[estimates of unwarranted racial disparity] may be the result of a race effect, but it may also stem from numerous

other factors that were not controlled" (p. 109). Further, Wilbanks lists several of these commonly omitted factors, such as: "the degree of premeditation, strength of evidence, willingness to plead guilty, willingness to testify against others, type of counsel, and prior record" (p. 110).

This argument suggests that studies that control for more variables and include controls for factors such as type of attorney, defendant socioeconomic status, method of disposition, and so forth produce systematically smaller estimates of unwarranted racial/ethnic disparities than other studies. Wilbanks' argument may have merit as some research indicates that several of the factors mentioned by Wilbanks do have meaningful relationships to sentence severity *and* are correlated with race. The effect of pleading guilty is a prime example. It can be stated unambiguously than defendants who plead guilty receive less severe sentences than defendants who are convicted by trial (e.g., see Albonetti 1990, 1997; Bushway and Piehl 2001; Dixon 1995; Engen and Steen 2000; Spohn 2000b). Interestingly, there is evidence to suggest that minorities, particularly African-Americans, may be less likely to plead guilty (Albonetti 1990; LaFree 1985; Petersilia 1983; Welch et al. 1985).

Pre-trial release status is another prime example of a variable associated with both outcome and race/ethnicity. Pre-trial release status has been found to have a strong positive relationship to sentence severity (Chiricos and Bales 1991; Hagan et al. 1980; Spohn and DeLone 2000; Spohn and Cederblom 1991).[6] Likewise, minorities,

[6] It should be noted that the relationship between pre-trial release and sentence severity may be produced by selection bias. It seems likely that serious offenders (i.e., defendants facing grave offenses and with long criminal histories) as well as offenders in cases where the evidence is strong, presumably would be less likely to be released during pre-trial and these types of offenders/cases would also be most likely to be punished severely. Thus, unless all of the factors affecting pre-trial release are included in the analyses of sentence severity (e.g., strength of evidence), the influence of pre-trial release would be biased by unmeasured factors (e.g., strength of evidence).

particularly African-Americans, have been found to be less likely to gain pre-trial release than whites (Spohn and DeLone 2000; Holmes et al. 1996). Type of attorney appears to have a less consistent relationship to sentencing severity than method of disposition or pre-trial release status; however, some research has found type of defense counsel to be related to sentencing severity. In particular, retaining a private attorney has been found to be associated with less punitive sentences (Chiricos and Bales 1991; Holmes et al. 1996; Spohn and Cederblom 1991; Unnever, 1982). Importantly for the current research, African-Americans have been found to be less likely to retain private attorneys (Holmes et al. 1996; Spohn, Gruhl, and Welch 1981-1982). Moreover, as previously mentioned, defendant SES has been found to have a negative relationship to sentence severity and SES is also negatively related to defendant minority status.

These findings suggest that studies controlling for type of counsel, method of disposition, defendant SES, and pre-trial status are less likely to confound the influence of race with the influence of these factors. Therefore, studies that employ such control variables may yield results that differ systematically from those studies that do not take into account these factors. In order to test this expectation, and as a test of Wilbanks' assertion that apparent race effects are in actuality due to model misspecification stemming from omitted variable bias, we coded separate indicators reflecting whether each effect size was produced by an analysis that included controls for type of attorney (private/retained or public/appointed), defendant socioeconomic status (employment status, income, and education), method of case disposition (plea/trial, plea bargained/not plea bargained), and pre-trial release status (released/not released).

Another methodological issue of great discussion in this body of literature revolves around the issue of sample selection bias. Several authors (e.g., Klepper et al. 1983; Peterson and Hagan 1984; Woolredge 1998; Zatz and Hagan 1985) have argued that examining only the sentencing stage may bias estimates of unwarranted disparity, as cases reaching this stage are not representative of cases subject to criminal sanctioning. That is, these cases are a biased sub-sample of cases eligible for punishment. In turn, utilizing these biased samples produces biased estimates of unwarranted disparity. In fact, Klepper et al. (1983) argue that "sample selection bias is likely to cause all the studies to underestimate the magnitude of discrimination in sentencing decisions" (p. 101). If Klepper and colleagues are correct than studies that attempt to account for possible selection bias should produce systematically larger estimates of unwarranted racial/ethnic disparity than those that do not. Thus, we distinguished studies that include such controls from those that do not.

We also coded moderator variables describing offense type, as there is evidence that the magnitude of unwarranted disparity varies by type of primary offense. For example, relatively large unwarranted disparities have been found in drug offenses, whereas smaller disparities have been found in studies analyzing violent offenses (see previous chapter). Thus, type of offense may be an important factor in explaining variation in unwarranted disparity.

Another interesting issue concerns the possibility of publication bias in previous reviews. All of the major reviews of the research have focused primarily on published research, leaving these reviews susceptible to publication bias. To test whether published studies are a biased sub-sample of all studies, we coded publication status (published

versus unpublished). Where "published" studies were defined as research published in journals, books, or book chapters, and all other studies have been coded as "unpublished."

We also coded features of the sentencing context such as presence of structured sentencing and region of jurisdiction. There is some evidence indicating that the implementation of structured sentencing mechanisms, particularly determinate sentencing and sentencing guidelines have reduced unwarranted racial/ethnic disparities (Klein et al. 1990; Miethe and Moore 1987; Tonry 1996). Furthermore, Kleck (1981) and Chiricos and Crawford (1995) have found unwarranted racial disparity was greater in Southern jurisdictions than other regions of the United States. Lastly, given the large concentration of Latinos in the Southwestern United States, the way in which Latinos, in comparison to non-Hispanic, are sentenced may differ from sentencing patterns in other regions of the U.S.[7]

The coded moderator variables described above have been organized into four categories describing sample characteristics, research methodology, and sentencing context. See Table 3 for a complete list and description of all coded moderator variables.

[7] Thanks to Terance Miethe for suggesting this moderator variable.

Table 3. Coded Moderator Variables

Moderator Variable	Values
Sample Characteristics	
Proportion Female	Continuous (0 to 1)
Proportion Black/Non-White	Continuous (0 to 1)
Proportion Latino	Continuous (0 to 1)
Proportion Asian	Continuous (0 to 1)
Proportion Native American	Continuous (0 to 1)
Mean Age	Continuous
Research Methodology	
Total Number of Variables	Continuous
Number of Criminal History Variables	Continuous
Number of Offense Seriousness Variables	Continuous
Number of Control Variables	Continuous
Nature of Criminal History Measure	0=Single Dichotomy (e.g. no prior convictions vs. some convictions); 1=Multiple Dichotomies/ Ordinal or Higher
Type of Criminal History Measure	1=Prior Arrest/Charge; 2=Prior Conviction; 3=Prior Incarceration; 4=Composite or Multiple Measures; 5=Other; 6=Unspecified
Type of Offense Seriousness Measure	1=Common Law Offense Types; 2=Offense Rating; 3=Common Law and Offense Rating
Controls for Type of Defense Counsel	0=No; 1=Yes
Controls for Method of Case Disposition	0=No; 1=Yes
Controls for Selection Bias	0=No; 1=Yes
Controls for Pre-trial Release Status	0=No; 1=Yes
Controls for Victim Injury	0=No; 1=Yes
Controls for Possession/Use of Weapon	0=No; 1=Yes
Controls for Defendant SES	0=No; 1=Yes
Questionable Analysis	0=No; 1=Yes
Publication Status	0=No; 1=Yes

Table 3. (cont.) Coded Moderator Variables

Precision of Race Measure	0=Non-Whites (i.e., mixes African-Americans with other minorities); 1=African-Americans Only
Sentencing Context	
Jurisdiction Type	1=City/County; 2=State; 3=Federal; 4=Other
Structured Sentencing	0=No; 1=Yes
Before 1980	0=No; 1=Yes
Southern Jurisdiction	0=No; 1=Yes
Southwestern Jurisdiction	0=No; 1=Yes

Coding Procedures and Quality Control

In order to ensure reliability of coding, we coded each study twice, once immediately after the study was retrieved and a second time after the search for eligible studies had been completed. Any discrepancies between codings were resolved in accordance to the coding manual. Copies of the coding forms utilized in this research are included in Appendix A.

Analytic Strategy

The analysis of effect sizes proceeds in two steps. First, we present a descriptive analysis of effect sizes. This descriptive analysis is analogous to descriptive statistics commonly reported in primary studies. Second, the coded effect sizes are analyzed via meta-analytic analogs to analysis of variance (ANOVA) and multiple regression. These

analyses determine which moderator variables are associated with observed variability in effect size (i.e., unwarranted sentencing disparity).

Descriptive Analysis

Using the moderator variables coded from each eligible study, the existing research will be described in regards to sample, contextual, and methodological characteristics. This description of the research yields a systematic audit of the extant research, which is necessary not only to characterize what has been accomplished but, perhaps more importantly, to reveal gaps in the research. In particular, the descriptive analysis presents descriptive statistics and graphics depicting the distribution of the effect sizes and moderator variables.

Effect Size Analysis

We utilize the statistical approach outlined by Lipsey and Wilson (2001) and Wang and Bushman (1999). In all analyses each effect size is weighted. Preliminary, each effect size is weighted by its inverse variance. The inverse variance weighting method has been shown to be the most efficient and accurate approach to incorporating the differential precision of effect sizes based on studies of varying sample size (Hedges 1982; 1994). The variance of the logged odds-ratio is defined as:

$$v_{lor} = \frac{1}{n_{m1}} + \frac{1}{n_{m0}} + \frac{1}{n_{w1}} + \frac{1}{n_{w0}} \quad (4)$$

where n_{m1} is the number of minority defendants (or cases involving minorities) who experience the event of interest, n_{m0} is the number of minority defendants who do not experience the event of interest, and n_{w1} and n_{w0} are defined similarly for white defendants. Because the terms in the denominators of equation 5 are not typically reported and many studies fail to report the standard errors associated with the odds

ratios, we estimated each of the terms by utilizing the reported odds ratio, marginal probability of punishment, and number of defendants of each race/ethnicity in the sample, as follows:

$$n_{m1} = \frac{-b - \sqrt{b^2 - 4(a)(c)}}{2(a)} \tag{5}$$

where $a = OR - 1$; $b = n_c - n_{r2} - OR(n_c) - OR(n_r)$, $c = OR(n_c)(n_r)$, OR is the odds ratio reported by the primary authors, n_c is the marginal number of defendants in the first column of a 2 x 2 contingency table based on the descriptive statistics reported in each study, n_r is marginal the number of defendants in the first row of the same contingency table, and n_{r2} is the marginal number of defendants in the second row of the contingency table. Once n_{m1} has been estimated, the rest of terms are estimated by subtraction (see Appendix B for an example of this process). We tested the accuracy of this estimation procedure by comparing standard errors obtained from the estimation process to standard errors reported by primary authors. In particular, we correlated the estimated standard errors with those reported by primary authors. The correlation coefficient between these standard errors was 0.90—indicating this estimation procedure was quite accurate.

The variance of the standardized mean difference effect size is defined as:

$$v_d = \frac{n_w + n_m}{n_w n_m} + \frac{ES_d^2}{2(n_w + n_m)} \tag{6}$$

where the terms are defined as above. Thus, the weight used for analysis is simply the inverse of these variances, or:

$$w = \frac{1}{v} \tag{7}$$

These weights imply a fixed-effects model. Fixed-effects models indicate that the only source of variability among the effect sizes is sampling error. That is, fixed-effects models assume that the distribution of effect sizes is homogenous.

This assumption of effect size homogeneity was tested in each analysis using the *Q* statistic as described by Lipsey and Wilson (2001:115). Given that a homogeneous distribution would display no more variability than that expected from sampling error alone, a statistical test of the homogeneity (*Q*) assumption is:

$$Q = \sum_{i=1}^{k} w_i (ES_i - \overline{ES})^2 \tag{8}$$

where *k* is the total number of effect sizes, and *Q* follows a chi-square distribution with *k* – 1 degrees of freedom. If *Q* exceeds the critical value of the chi-square distribution with *k* – 1 degrees of freedom, then the null hypothesis is rejected. Such a finding indicates that sources of variability beyond sampling error exist and, therefore, each effect size does not estimate the same population mean.

The vast majority of the homogeneity analyses employed in this analysis indicated that the distribution of effect sizes exhibited more variation than would be expected by only sampling error. Further, even after taking into account our coded moderator variables into account, the residual variation continued by exhibit more variability than that expected by sampling error. Therefore, a random effects component was added to the weights to capture unmeasured (random) differences between studies, as follows:

$$v^* = v_i + v_\theta \tag{9}$$

where v_i is the sampling error variance and v_θ is the random effects variance. The random (mixed) effects variance component captures other sources of variability above and beyond sampling error. This approach is more conservative than the fixed effects approach in that it produces larger confidence intervals around the mean effect sizes (for a discussion of the random effects model see Lipsey and Wilson 2001; Overton 1998; Raudenbush 1994). In fact, given that the studies reviewed in the present meta-analysis were not randomly selected,[8] the random effects approach probably overestimates the actual variability among studies and as a consequence creates confidence intervals that are too large (Overton 1998). In order to avoid being overly conservative, we interpret moderator relationships that are marginally statistically significant (i.e., $p < 0.10$) as being meaningful.

The primary analytic tools employed for determining which moderator variables are statistically associated with effect size were the meta-analytic analogs to ANOVA and weighted multiple regression.[9] As a first stage in the data analysis, we conducted a series of bivariate analyses, which tests whether each moderator variable has a statistically significant bivariate relationship with effect size. The relationship between effect size and categorical or ordinal variables were analyzed via meta-analytic ANOVA, whereas the bivariate relationship between effect size and continuous measures were analyzed using weighted mixed-effects (i.e., fixed slope and random intercept) simple regression. The second stage of the data analysis regresses the dependent variable (effect size) on those moderator variables that displayed signs of meaningful bivariate relationships to

[8] The current meta-analysis aims to be comprehensive; thus, the studies included in this research were not randomly selected.

[9] These analyses utilized macro programs created by David Wilson. As of this writing, David Wilson has made these macro programs available to the public at: http://mason.gmu.edu/~dwilsonb/ma.html

effect size using a full-information maximum likelihood multivariate mixed-effects model (see Lipsey and Wilson 2001; Raudenbush 1994). These analyses are repeated for each of the sentencing outcome measures and all analyses are based on the appropriate weighting method (i.e., fixed or random effects models). Lastly, separate parallel analyses are conducted for sentencing outcomes relating to African-American and Hispanics.

Limitations of Research

While we believe that the current research is a marked improvement over existing syntheses, it has several weaknesses that should be acknowledged. In our opinion, the most important limitation of this research is its focus on the direct influence of race/ethnicity on sentencing outcomes. There is considerable evidence indicating that race/ethnicity has important indirect influences (e.g., see LaFree 1985; Lizotte 1978); unfortunately, this work is too diverse and scattered to be meaningfully synthesized quantitatively. Furthermore, many studies do not discuss whether the indirect effects of race/ethnicity were assessed. This leads to ambiguity concerning whether there were no meaningful indirect effects or did the author(s) simply fail to test for these effects. For many of the same reasons, the present research does not focus on the interactional effects of race/ethnicity on sentencing outcomes.

Another limitation of this research is that institutionalized racism is not addressed by this research. This research focuses on whether sentencing policies are applied in a race/ethnicity neutral fashion; the question of whether these policies are written in a race

neutral manner is outside the purview of this research. However, it is also important to question whether sentencing laws are themselves racially biased.

Furthermore, the method of meta-analysis has typically been criticized on several recurring issues. The first issue is what we refer to as the "apples and oranges" problem. In essence, this criticism makes the point that some meta-analyses include studies that operationalize the dependent variable in too many, disparate manners to be meaningfully combined. The second criticism is that meta-analyses mix studies of different methodological rigor.

We believe that this meta-analysis does not suffer from these weaknesses. First, because the dependent variable in sentencing research has been operationalized in relatively few distinct manners, and because we are analyzing each of the major outcomes separately, we do not believe that the issue of comparing apples to oranges is problematic. Second, while there is undeniably a great deal of methodological variation between sentencing studies, the criterion that all studies control for offense seriousness and prior criminal conduct seems to be a reasonable lower limit. Moreover, this research attempts to capture variation in methodological rigor with the coded moderator variables.

An additional potential threat to the findings of the current research is that it relies on a body of primary research that is replete with potential methodological flaws. Many of the studies included in this meta-analysis utilize questionable statistical controls, fail to include controls for important third factors (i.e., variables related to both sentence severity and race/ethnicity, such as defendant SES), and arguably commit other specification errors. Whether these methodological flaws actually lead to biased estimates of unwarranted sentence disparity is an empirical issue that this meta-analysis

attempts to address. Rather than establish more strenuous, but arbitrary, inclusion criteria, we have attempted to code relevant methodological features. However, to the extent that *uncoded* methodological features are related to estimates of unwarranted racial/ethnic disparity, the results of the present study may be suspect. Stated differently, many of the studies included in the present research are methodologically flawed; however, if the coded study features capture relevant methodological variation then the inclusion of these studies is not problematic. On the other hand, if the coded study features are inadequate in capturing methodological variation, then the present meta-analysis will yield questionable results.

CHAPTER 4. RESULTS

Descriptive Analysis

The search strategy described in the previous chapter uncovered 336 potentially eligible studies; i.e., studies that could not be ruled ineligible from a review of the study's abstract. A full version of each of these studies was retrieved with the exception of five studies that we was unable to locate. After screening the full version of each of these 331 studies, we determined that 184 studies (55%) met the eligibility criteria, and 147 studies (45%) were ineligible for various reasons (see Table 4).

Table 4. Summary of Final Eligibility Status

Final Eligibility Status	N (%)
Total Number of Studies Identified	336 (100%)
Ineligible	147 (44%)
--No Simultaneous Control of Offense serious/Prior Record	41 (28%)
--No Empirical Analysis	40 (27%)
--Did Not Measure Direct Effect of Race/Ethnicity	25 (17%)
--No Sentencing Outcome	39 (24%)
--Other	2 (1%)
Eligible	184 (54%)
--But Statistically Dependent	80 (43%)
--Same Data, Different Outcome	9 (5%)
--Uncodeable[a]	19 (10%)
Unable to Retrieve	5 (2%)
Eligible, Statistically Independent, and Codeable	76

a. Seven studies were uncodeable because no standard deviations were reported, seven other studies were uncodeable because of the type of analysis utilized by the primary authors, and five studies were uncodeable because the numerical values of parameters were not reported.

Table 4 shows that ineligible studies failed to meet the inclusion criteria for one of three primary reasons. Studies were ruled ineligible were because they did not: 1) simultaneously control for seriousness of current offense and prior criminal history of the defendant; 2) examine any of the five specific sentencing outcomes encompassed by this

meta-analysis (as specified in the last chapter); or, 3) conduct any empirical analyses of individual-level court outcomes. Approximately 80% of ineligible studies were declared ineligible for one of these three reasons. A smaller percentage of studies (17%) were ruled ineligible because they did not include a measure of race/ethnicity or did not measure the direct effect of race/ethnicity on a sentencing outcome.

After determining each study's final eligibility status, eligible studies were cross-checked against one another to ensure that each study was statistically independent; that is, no two analyses of the same data set *and the same sentencing outcome* were allowed to be included in this meta-analysis. This cross-checking procedure indicated that many of the eligible studies were linked to one another. In fact, 80 of the 184 eligible studies (43%) analyzed the same data and same sentencing outcome as another study included in this meta-analysis, hereafter these studies will be referred to as "dependent studies." Thus, 104 studies of the eligible studies were statistically independent. This number is further reduced by the fact that nine studies analyzed the same data set, but analyzed a different sentencing outcome as another study; hereafter these studies are referred to as "related studies." Studies analyzing the same data *but different sentencing outcomes* were collapsed into one study with multiple outcomes for the purposes of this meta-analysis.

Nineteen of the remaining eligible, independent studies did not report enough information to calculate an effect size or the analytic technique employed was unsuitable for effect size coding. There were three primary reasons that prohibited effect size calculations: 1) the author(s) did not report the standard deviation of the dependent variable (or sufficient information to estimate its standard deviation); 2) the author(s) did

not report numerical values of parameter estimates (e.g., the authors did not report regression coefficients); or, 3) the type of analysis was uncodeable (e.g., structural equation modeling, stepwise regression, log-linear analysis).

In all, the total number of eligible, independent, and codeable studies is 76—these are the studies included in the following analysis, hereafter referred to as "coded studies." These 76 coded studies actually capture the results from 85 studies as nine related studies were collapsed into the 76 coded studies. All 85 coded and related studies are indicated in the bibliography by an asterisk (*). The full citations for the 79 dependent studies are included in Appendix C. Likewise, the full citations for the 19 eligible but uncodeable studies are listed in Appendix D. Appendix E contains the full citations for each of the 147 ineligible studies, listed by ineligibility reason.

Table 5 displays information regarding publication type and year of publication. Approximately half of eligible, independent studies were published as journal articles (49%), and another 14% of studies were published as books or book chapters. A considerable proportion of eligible independent studies, however, were unpublished (37%). The large percentage of unpublished studies included in the present meta-analysis reduces the possibility of publication bias distorting the findings of this cumulative body of research. Furthermore, half of these unpublished studies were doctoral dissertations (50%), which generally displayed a level of methodological and analytical rigor comparable to published studies.

Interestingly, while the question of the racial/ethnic neutrality of the sentencing in the United States has been a long-standing research focus, Table 5 shows that the bulk of studies meeting the inclusion criteria for this meta-analysis were published in the 1990s

(45%) or 1980s (32%). A smaller percentage of studies (9%) were published in the 1970s or since 2000 (11%), and only three of the studies (4%) included in this analysis were published in the 1960s. Thus, the eligibility criteria for this study systematically excluded earlier studies of unwarranted disparity in sentencing outcomes, as these early studies tended to lack the requisite methodological rigor necessary for inclusion.

Table 5. Publication Type and Year

Publication Characteristic	N (%)
Publication Type	
Book	7 (9%)
Book Chapter	4 (5%)
Journal Article	37 (49%)
Unpublished	28 (37%)
Publication Year	
1960-1969	3 (4%)
1970-1979	7 (9%)
1980-1989	24 (32%)
1990-1999	34 (45%)
2000-2002	8 (11%)

The number of effect sizes available for this analysis is considerably greater than the total number of studies. In fact, the 76 coded studies produced 122 independent sentencing contexts, as roughly 30% of the eligible independent studies reported multiple sentencing contexts (i.e., time periods and places). That is, the 76 independent studies reported sentencing outcomes from 122 sentencing contexts; these 122 sentence contexts serve as the primary unit of analysis in this meta-analysis. Furthermore, because many sentencing contexts reported analyses of multiple sentencing outcomes and/or multiple racial/ethnic contrasts a total of 430 effect sizes were coded. The bulk of these effect sizes compared sentencing outcomes of African-Americans to those of whites (66%), 25% of effect sizes contrasted sentencing outcomes for Latinos to whites, whereas only

4% and 5%, respectively, concerned sentencing outcomes of Asians and Native Americans in comparison to whites.

Interestingly, Table 6 reveals that the most common type of sentencing outcome analyzed in the coded sentencing contexts related solely to imprisonment decisions (37%). Another sizeable proportion of sentencing contexts (30%) analyzed sentencing data involving multiple types of sentencing outcomes (e.g., imprisonment and sentence length decisions). A sizeable but smaller proportion of analyses considered only sentencing outcomes related to length of incarceration sentence. Still other sentencing contexts utilized measures of sentence outcomes that simultaneously combined imprisonment and sentence length decisions. Discretionary leniency and discretionary harshness were the least common types of sentencing outcomes.

Table 6. Type of Sentencing Outcome Analyzed

Sentencing Outcome	k^a (%)
Imprisonment Decision	45 (37%)
Length of Incarcerative Sentence	20 (16%)
Simultaneous Analysis of Imprisonment/Sentence Length	14 (11%)
Discretionary Lenience	3 (2%)
Discretionary Harshness	3 (2%)
Mixture of Above Categories	37 (30%)

a. *k* = number of effect sizes with non-missing values.

Table 7 reports descriptive statistics on the contextual characteristics of the 122 independent sentencing contexts included in this meta-analysis. A little less than half of the sentencing contexts (44%) analyzed data from cities or counties. Another 41% of contexts analyzed state level data (from a single state), 13% of contexts analyzed data from Federal courts, and the remaining 2% of contexts were classified as "other"(e.g., pooled court data from multi-states). In 31% of the sentencing contexts, some form of

structured sentencing was employed. Most often, these contexts utilized sentencing guidelines, while a considerably smaller number of contexts applied determinate sentencing.

Table 7. Sentencing Context Characteristics

Contextual Characteristic	*k*[a] (%)
Jurisdiction Type	
City/County	54 (44%)
State	50 (41%)
Federal	15 (13%)
Other	3 (2%)
Structured Sentencing	
Yes, presumptive sentencing guidelines	26 (21%)
Yes, voluntary sentencing guidelines	1 (1%)
Yes, determinate sentencing	11 (9%)
No	84 (69%)
Before 1980[b]	
Yes	58 (47%)
No	63 (52%)
Unknown	1 (1%)
After Drug War	
Yes	28 (23%)
No	94 (77%)
Southern[c]	
Yes	25 (21%)
No	82 (67%)
Not Applicable (Federal, Mixture of States, or Unknown)	15 (12%)
Southwestern[d]	
Yes	21 (17%)
No	85 (70%)
Not Applicable (Federal, Mixture of States, or Unknown)	16 (13%)

a. *k* = number of effect sizes with non-missing values.
b. Categorization is based on mid-point of data series; i.e., contexts whose data mid-point is prior to 1980 are classified as "1."
c. Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia.
d. Arizona, California, Colorado, Nevada, New Mexico, and Texas

The data analyzed in the coded studies date as far back as 1929 and as recently as 2000. Categorizing each sentencing context by the midpoint of the data series, it is

apparent that few analyses analyzed data with a mid-point prior to 1970; only 7% of contexts analyzed data with a mid-point prior to 1970, 42% of contexts analyzed data whose mid-point occurred in the 1970s, 34% analyzed data collected in the 1980s, and 17% analyzed data collected since the 1980s. Moreover, 47% of the contexts analyzed data prior to the sentencing reform era (i.e., prior to 1980). Twenty-three percent of sentencing contexts analyzed data concerning sentences imposed after the commencement of the war on drugs (operationalized as 1987 and afterwards).[10] Thus, while a disproportionate number of studies were published in the 1980s and 1990s (77%), only about half of the included studies analyzed data from the 1980s or later.

Geographically, the 122 coded sentencing contexts analyzed sentencing practices in the majority of States. Twenty-one percent of the sentencing contexts analyzed data collected from the 11 former Confederate states and 17% contexts involved data from Southwestern states (i.e., Arizona, California, Colorado, Nevada, New Mexico, and Texas). As Table 8 shows, several states' sentencing practices were analyzed repeated. In particular, sentencing practices in California, Pennsylvania, New York, and Florida were subjected to numerous empirical analyses–in all, nearly 40% of the sentencing contexts in this meta-analysis concern sentencing practices in these four states. An additional 11% of sentencing contexts concern sentencing practices in the Federal courts.

[10] Thanks to Gary LaFree for suggesting this moderator variable.

Table 8. Sentencing Contexts by Jurisdiction

Jurisdiction	*k*[a] (%)
Alabama	2 (2%)
Alaska	1 (1%)
Arizona	3 (2%)
California	9 (7%)
Colorado	1 (1%)
Connecticut	1 (1%)
Florida	11 (9%)
Georgia	1 (1%)
Illinois	1 (1%)
Iowa	3 (2%)
Kentucky	2 (2%)
Louisiana	1 (1%)
Maryland	1 (1%)
Massachusetts	2 (2%)
Michigan	2 (2%)
Minnesota	5 (4%)
Missouri	1 (1%)
New Jersey	1 (1%)
New Mexico	1 (1%)
New York	18 (15%)
North Carolina	2 (2%)
Ohio	4 (3%)
Oklahoma	4 (3%)
Pennsylvania	10 (8%)
South Dakota	1 (1%)
Texas	7 (6%)
Virginia	1 (1%)
Washington, D.C.	1 (1%)
Washington	5 (4%)
Wisconsin	3 (2%)
More than one State	1 (1%)
Unknown	3 (2%)
Federal	13 (11%)

a. *k* = number of effect sizes with non-missing values.

Methodologically, the analyses employed in the 122 sentencing contexts included a sizeable number of variables. On average, approximately 11 variables were utilized in the analyses (see Table 9). Most of these variables were categorized as control variables (i.e., not related to measuring defendant's race/ethnicity, offense seriousness, or criminal history).

Table 9. Descriptive Statistics on Methodological Variables

Methodological Variable	Mean (SD)	Minimum	Maximum	k^a
Total Variables	10.99 (6.45)	2	39	121
Criminal History Variables	1.71 (1.61)	0	12	121
Offense Seriousness Variables	2.63 (1.86)	0	11	121
Control Variables	5.75 (4.95)	0	29	121

a. *k* = number of effect sizes with non-missing values.

Further, as Table 10 illustrates that the most common type of control variables employed in these coded analyses were related to method of disposition (plea vs. trial; 57%) or the SES of the defendant (41% of analyses controlled for defendant SES). Other relatively common control variables concerned presence of a weapon, pre-trial status of the defendant (released vs. in-custody), and type of defense counsel (private/retained or public/appointed). Few studies included controls for victim injury or utilized analytic techniques designed to reduce the possibility of selection bias.

Table 10. Descriptive Statistics Methodological Features

Methodological Feature	k^a (%)
Criminal History Level of Measure	
Single Dichotomy	25 (20%)
Multiple Dichotomies/Ordinal or Higher	97 (80%)
Type of Criminal History Measure	
Arrest/Charges Filed	4 (3%)
Conviction	48 (39%)
Incarceration	14 (11%)
Multiple	51 (42%)
Other	1 (1%)
Unspecified	4 (3%)
Type of Offense Seriousness Measure	
Common Law	36 (30%)
Offense Rating	38 (31%)
Common Law and Offense Rating	48 (39%)
African-Americans Only (or Mixed with Races)	
Yes	73 (72%)
No	28 (28%)

Table 10 (cont.) Descriptive Statistics Methodological Feature

Controls for Method of Disposition	
Yes	69 (57%)
No	53 (43%)
Controls for Defendant SES	
Yes	50 (41%)
No	72 (59%)
Controls for Type of Counsel	
Yes	37 (30%)
No	85 (70%)
Controls for Possession/Use of a Weapon	
Yes	37 (30%)
No	85 (70%)
Controls for Pre-trial Release Status	
Yes	36 (30%)
No	86 (70%)
Controls for Selection Bias	
Yes	15 (12%)
No	101 (88%)
Controls for Victim Injury	
Yes	12 (10%)
No	110 (90%)
Questionable Analysis	
Yes	34 (28%)
No	88 (72%)
Unpublished	
Yes	38 (31%)
No	84 (69%)

a. *k* = number of effect sizes with non-missing values.

In regards to offense seriousness, relatively few variables, on average approximately three variables, were utilized to capture variability on this important factor. In approximately 40% of the coded analyses, seriousness of current offense was measured utilizing a combination of offense severity ratings (e.g., sentencing guideline scores) and common law offense types (e.g., drug offenses, violent/person offenses). The remainder of the coded analyses was evenly split between those that measured offense seriousness utilizing only common law offense types (30%) or only offense severity ratings (31%).

Table 9 reveals that even fewer variables, on average less than two variables, were employed to measure defendant's prior criminal history. The overwhelming majority (79%) of the coded analyses measured defendants' criminal history using either a non-dichotomous criminal history measure (i.e., measures scaled at the ordinal level or higher) or multiple dichotomous variables (see Table 10). Most often, the criminal history measures utilized in these analyses (42%) captured multiple indicators (e.g., prior arrests, convictions, incarcerations) of criminal behavior. A little less common were measures of criminal history that relied only on information regarding the number of prior convictions (39%). Relatively few studies measured criminal history using only information concerning defendants' prior history of incarceration or arrest.

Overall, the methodological rigor of this body of research appears to be increasing. One of the strongest indicators of the increasing methodological rigor evident in this body of research is demonstrated by tracking changes in the above methodological variables over time. For example, studies published prior to 1980 averaged approximately three control variables, whereas since 1980 the mean number of control variables has increased to six control variables. Similarly, 36% of studies made available prior to 1980 utilized questionable analytical techniques, in comparison 25% of studies made available since 1980 employed such techniques. Another strong indicator of the increasing methodological rigor of this body of research is found by comparing the manner in which criminal history is most commonly measured in the analyses included in the present synthesis to that most commonly employed in earlier syntheses of this body of research. For example, in this meta-analysis 79% of coded analyses used relatively precise measures of criminal history (i.e., multiple dichotomous measures or measures

scaled at the ordinal level or higher); in stark contrast, less than 40% of the studies included in Kleck's review employed such rigorous measures.

In spite of these recent methodological improvements, a significant proportion of analyses were of questionable analytical rigor or utilized questionable measures of race. In fact, 27% of the coded sentencing contexts examined data using techniques that are generally regarded as flawed, such as analyzing a dichotomous outcome using OLS regression or utilizing OLS regression with an arbitrary, non-interval scale dependent variable. Similarly, 35% of coded analyses employed questionable measures of race. Typically, these studies measured race or minority status by lumping (primarily) African-American defendants with a smaller number of defendants from other racial/ethnic minority group. Implicitly, these studies assume that the effect of minority status does not vary by specific minority groups. As a result, these studies potentially introduce an additional source of measurement error.

Not surprisingly, the samples analyzed in this meta-analysis were comprised predominantly of young males, and minorities (see Table 11). The mean sample age was 28 years old and males comprised at least 80% of most samples. On average, African-Americans comprised 43% of samples in contexts comparing sentencing outcomes of African-Americans to whites, when Latinos' sentencing outcomes were compare to whites, Latinos represented 23% of these samples.

Table 11. Sample Characteristics

Sample Characteristic	Mean (SD)	Minimum	Maximum	k^a
Age	28.13 (3.68)	16.98	34.78	52
Proportion Female	0.19 (0.29)	0.00	1.00	98
Proportion Black/Non-White	0.43 (0.22)	0.02	0.94	111
Proportion Latino	0.23 (0.20)	0.01	0.80	37
Proportion Native American	0.11 (0.11)	0.02	0.25	7
Proportion Asian	0.02 (0.01)	0.01	0.03	5

a. k = number of effect sizes with non-missing values.

Summary of Descriptive Analysis

The descriptive analysis of the studies included in this meta-analysis reveals several important points. First, the current research appears to be the most comprehensive review of the research in this area, as this meta-analysis includes the results from 85 published and unpublished studies—a number greater than any of the major existing reviews. Further, this meta-analysis includes analyses of unwarranted racial disparity from 29 of the 50 states, as well as analyses of sentencing practices in the Federal courts and Washington D.C. The descriptive analysis also reveals that while there over 180 studies meeting the eligibility criteria for inclusion in this meta-analysis, a large proportion of these studies used the same data or overlapping data as another study, or failed to report enough information to calculate an effect size.

Second, it is evident that the methodological rigor in this body of research has improved markedly over the past two decades. Analyses of unwarranted racial/ethnic disparity have included an increasing number of control variables (i.e., variables not measuring race/ethnicity, criminal history, or offense seriousness). Most commonly, these controls measure defendant socioeconomic status, type of defense counsel, and method of case disposition—all of which have been found to be related to race/ethnicity and severity of sentencing outcomes.

Third, most sentencing research continues to focus on comparing sentencing outcomes of African-Americans to whites. In fact, 95% of coded studies included contrasts between African-Americans and whites, and 66% of all effect sizes contrasted sentencing outcomes of African-Americans to those of whites. However, an increasing

number of studies are including empirical assessments of Latino sentencing outcomes. In all, 25% of all coded contrasts compared sentencing outcomes of Latinos to whites and almost all of this research was published since 1980. Little research concerns sentencing practices of Native Americans or Asians as only 5% and 4% of coded contrasts, respectively, compared sentencing outcomes between these minority groups and whites.

Effect Size Analysis

African-American/White Contrasts

One hundred sixteen of the 122 (95%) independent sentencing contexts compared sentencing outcomes of African-Americans to those of whites. In all, these 116 sentencing contexts produced 282 effect sizes. Fifteen of these 116 sentencing contexts analyzed data from the Federal court system—producing 24 effect sizes; whereas the remaining 101 sentencing contexts analyzed data from State (i.e., non-Federal) sentencing contexts yielding 258 effect sizes. Each effect size was transformed onto the odds ratio scale using the conversion factor described in Lipsey and Wilson (2001), as this metric was the most common. Preliminary examination of the coded effect sizes indicated that analyses of data from the Federal courts differed in several important ways from analyses of State court data. Therefore, parallel analyses of Federal court and State court data are conducted on African-American/white effect sizes.

The vast majority of coded effect sizes indicated that African-Americans were sentenced more harshly than whites. In regards to Federal sentencing contexts, 83% of these effect sizes coded indicated that African-Americans were sentencing more harshly than whites. Similarly, 77% of the effect sizes calculated from State court data indicated

that African-Americans were sentenced more harshly than whites. This is preliminary evidence that African-Americans in both Federal and State sentencing contexts are sentenced more harshly than whites; however, these effect sizes are not statistically independent and therefore this finding is only suggestive of unwarranted racial disparity. Furthermore, while these statistics are suggestive of unwarranted racial disparity in sentencing outcomes, they are not very helpful in determining the magnitude or variability in the magnitude of observed differences between African-Americans and whites.

As a first step toward more rigorously addressing the issue of unwarranted racial disparity disfavoring African-Americans, we first analyze the overall unwarranted disparity measure. As described in the previous chapter (see pg. 40-42), in contexts where only one sentencing outcome was examined, and hence only one effect size was computed, the overall effect size measure is simply the effect size from this one contrast. However, in contexts that analyzed more than one type of sentencing outcome, most commonly imprisonment decisions and decisions regarding length of incarcerative sentence, then the overall effect size measure is the weighted average of the effect sizes from these separate analyses. Likewise, in contexts that reported multiple African-American/white contrasts by disaggregating by type of offense, the overall effect size was calculated by averaging the effect sizes computed for each offense type.

It is important to keep in mind that all effect sizes are coded such that positive effect sizes indicated that the minority group of interest was sentenced more harshly than whites, and all effect sizes are reported on the odds ratio metric (as this metric was the most common). Thus, an odds ratio greater than 1 denotes that the minority group of

interest was sentenced more harshly than whites, independent of prior criminal history and current offense seriousness. An odds ratio of 1 indicates minorities and whites were sentenced with equal severity; whereas, an odds ratio less than 1 indicates that whites were sentenced more harshly than the minority group of interest.

This process of effect size coding yielded 116 independent effect sizes; 101 of which were State sentencing contexts and the remaining 15 were Federal sentencing contexts. Analyzing these independent effect sizes continues to indicate that African-Americans on average were sentenced more harshly than whites. In fact, the results from these analyses are similar to the preliminary analysis of the 282 non-independent effect sizes; 73% of the independent effect sizes from the Federal system indicated that African-Americans were sentenced more severely than whites and 33% of the total number of effect sizes were statistically greater than 1. By contrast, 27% of the independent Federal effect sizes found that whites were sentenced more harshly than African-Americans. Likewise, 76% of the effect sizes calculated from State data showed that African-Americans were sentenced more harshly than whites.

In regards to the distribution of effect sizes from State data, this distribution of odds ratio effect sizes ranged from a modestly large negative (i.e., odds ratio less than 1) effect of 0.40, signifying that whites in this particular sentencing context were sentenced more harshly than African-Americans, to large positive effect of 8.41 indicating that African-Americans were sentenced much more harshly than whites in another specific sentencing context. The *Q*-statistic, which tests the assumption that the only source of effect size variation is sampling error (i.e., fixed effects), indicates that for the present sample of effect sizes this assumption is not tenable ($Q[100] = 2091$, $p < 0.0001$). This

indicates that the present distribution is not estimating a common population effect size, and therefore the assumptions underlying the random effects model are more plausible. The random effects mean odds ratio was 1.28 with the 95 percent confidence interval ranging from a lower bound of 1.20 to an upper bound of 1.35 (see Table 12a).

Table 12a. African-American/White Contrasts by Type of Effect Size Analysis

	State Data			
		95% C.I.		
Type of Analysis	Mean Odds Ratio	Lower	Upper	k^a
Fixed Effects	1.24***	1.23	1.25	101
Random Effects	1.28***	1.20	1.35	101
Unweighted	1.32***	1.08	1.60	101

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$. These tests reject the null hypothesis of a mean odds ratio equal to 1; i.e., rejects hypothesis of equality of sentencing severity between African-Americans and whites.

A more intuitive sense of this effect size can be gained by transforming this effect size into percentage. If we assume that 50% of whites were punished (e.g., incarcerated), then this overall mean odds ratio translates into a punishment rate of approximately 56% for African-Americans. This translation is for heuristic purposes only, as the assumed 50% rate of punishment for whites was arbitrarily chosen. Further, because of the non-linearity of the odds ratio, assuming a 50% rate of punishment for whites maximizes the percentage difference in punishment severity between whites and African-Americans. That is, if we assumed any other punishment rate for whites, the percentage difference between whites and African-Americans would be smaller.

The mean odds ratio effect size from the 15 sentencing contexts analyzing Federal court data also indicated that African-Americans on average were sentenced more harshly than whites, independent of offense seriousness and offender criminal history (see Table

12b). The effect of race in this analyses also is highly variable ($Q[14] = 169$, $p < 0.0001$), indicating that this distribution of effect sizes is not estimating a common population mean effect size (i.e., fixed effects models are not tenable); therefore, a random effects model was used. The results from this model reveal that the mean effect size from analyses of Federal court data is somewhat smaller than that of State court data (1.15 vs. 1.28); that is, unwarranted racial disparity in sentencing outcomes disadvantaging African-Americans was somewhat larger in analyses of State courts than in Federal courts. In fact, the random effects mean odds ratio of Federal sentencing contexts is not statistically significant at conventional levels of significant ($p = 0.093$) as the 95% confidence interval extends below 1. Translating the Federal mean odds ratio effect size into percentages suggests that 53% of African-Americans would be punished if we assume that 50% of whites would be punished; clearly, the influence of race in the Federal courts is statistically small.

Table 12b. African-American/White Contrasts by Type of Effect Size Analysis

	Federal Data			
		95% C.I.		
Type of Analysis	Mean Odds Ratio	Lower	Upper	*k*
Fixed Effects	1.33***	1.29	1.37	15
Random Effects	1.15*	0.98	1.34	15
Unweighted	1.16	0.70	1.92	15

a. *k* = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$. These tests reject the null hypothesis of a mean odds ratio equal to 1; i.e., rejects hypothesis of equality of sentencing severity between African-Americans and whites.

Figure 1 displays a random sample of the distribution of the 101 odds ratio effect sizes reflecting overall unwarranted racial disparities from State data in a forest plot. Figure 2 contains the same information for the 15 analyses of Federal data. In these

forest plots, each sentencing context included in this synthesis is identified on the left by the study's author(s) (and where necessary the name of the sentencing context or time period is listed in parentheses), year of publication; on the right side of the figure, the odds ratio effect size from each context is represented by a diamond and the 95 percent confidence interval is represented by line extending from the diamond. Those effects sizes that do not cross the centerline, which represents an odds ratio of 1, are statistically significant at the 0.05 level of significance. At the very bottom of each figure, the overall mean random effects odds ratio effect size and confidence interval is displayed. The large number of effect sizes made it impossible to display the effect size distribution on one plot; therefore, we took a 35% random sample of the 101 effect sizes from State data. This forest plot of randomly selected effect sizes closely resembles the complete forest plot. (The complete distribution of effect sizes is displayed in Appendix F; in this figure the effect sizes are displayed over three pages.)

Figure 1. African-American/White Contrasts: Forest Plot—State Level Only

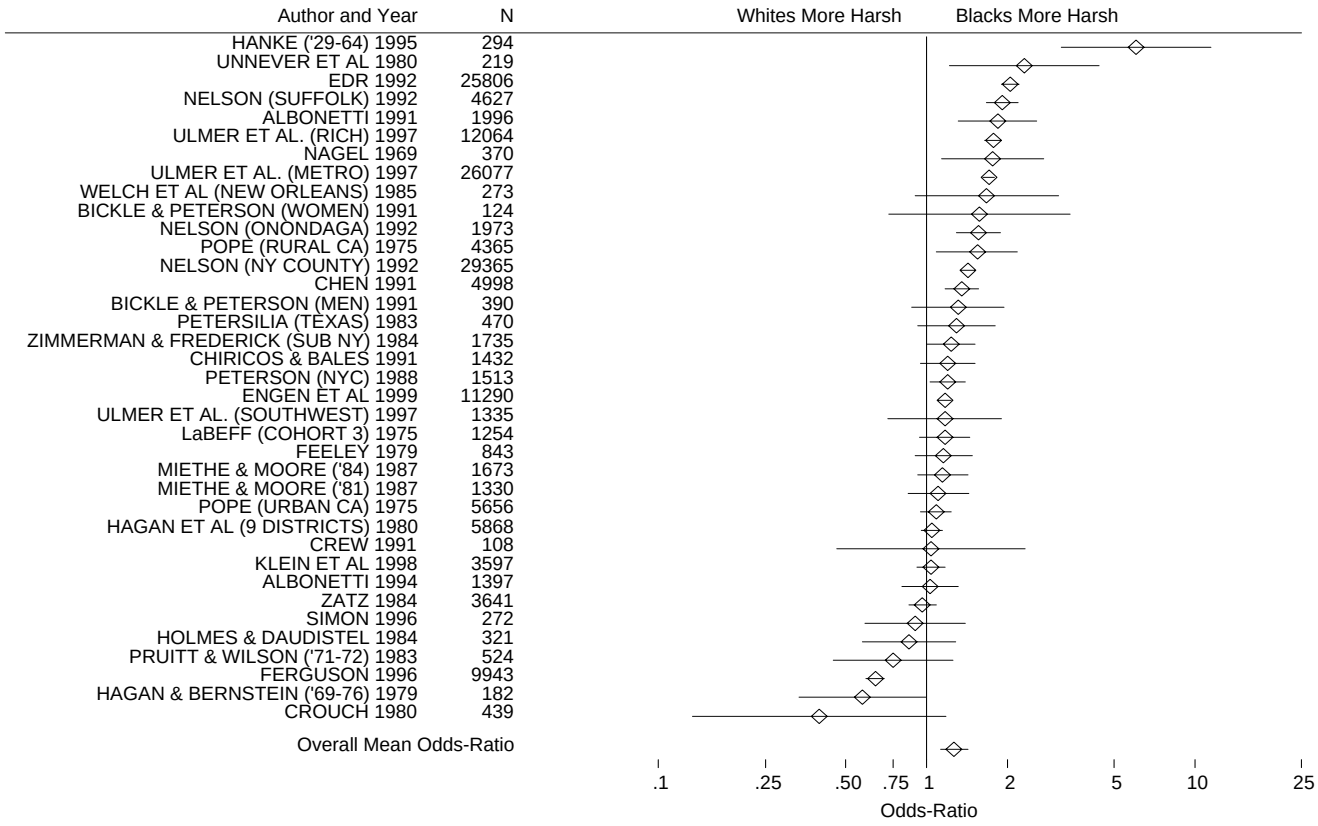


Figure 2. African-American/White Contrasts: Forest Plot—Federal Level Only

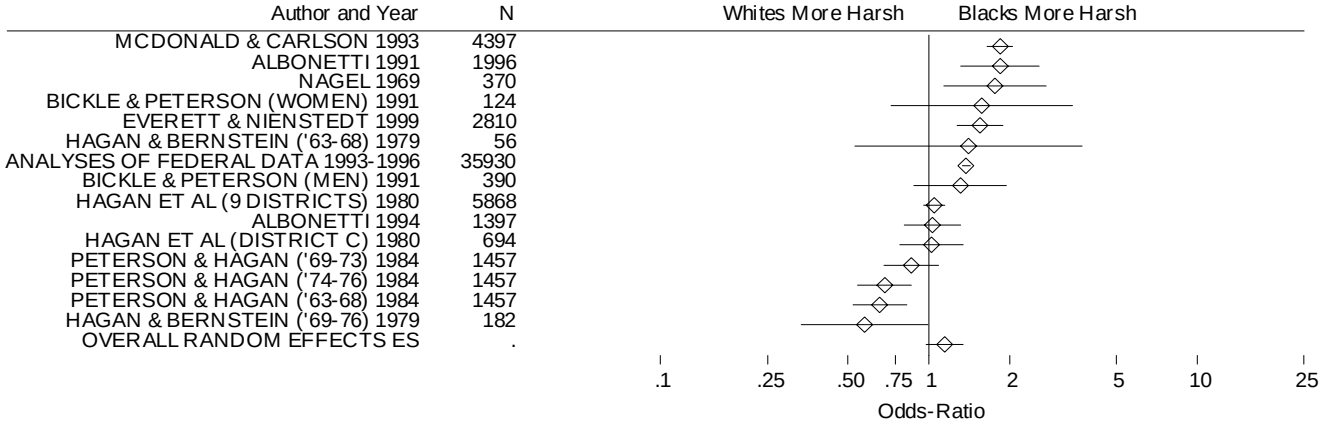


Figure 3 displays the two effect size distributions in a stem and leaf plot; in this stem and leaf plot the logged odds ratios are graphed.[11] From this figure it appears that the distribution of effect sizes from State data may be distorted by the presence of several effect sizes that are considerably larger than most other effect sizes (i.e., potential outliers). In order to remove the potentially distorting effects of these extreme effect sizes, we re-ran the above analyses after removing the most extreme 2.5% of the original effect size distribution from both ends of the distribution, yielding a trimmed 95% distribution of the original effect sizes. The overall mean random effects odds ratio of this trimmed distribution is 1.27 with a 95% confidence interval of 1.19 to 1.34, which is nearly identical to the same statistics for the original distribution of effect sizes.[12] Thus, the presence of potential outliers does not bias these results.

[11] Displaying logged odds ratios is more efficient than odds ratios, because of the non-linearity of odds ratios.
[12] we also re-ran the above analyses removing only the most extreme 2.5% of the upper end of the distribution. This analysis indicated that the random effects overall mean effect size is 1.25 with a confidence interval of 1.18 to 1.33.

Figure 3. African-American/White Contrasts: Stem and Leaf Plot			
State Data		Federal Data	
-9*	2		
-8*			
-7*			
-6*			
-5*		-5*	5
-4*	664	-4*	2
-3*		-3*	8
-2*	9643	-2*	
-1*	53100	-1*	5
-0*	5321	-0*	
0*	00000000244466788	0*	235
1*	01122345555566677889	1*	
2*	00015567	2*	7
3*	012223366	3*	24
4*	1145589	4*	45
5*	237899	5*	7
6*	457999	6*	11
7*	126		
8*	24		
9*			
10*	5		
11*	2		
12*			
13*			
14*			
15*	26		
16*			
17*			
18*	0		
19*			
20*			
21*	3		

The above findings, in agreement with our first hypothesis, indicate that African-Americans when sentenced in State courts were generally punished more harshly than whites, independent of offense seriousness and prior criminal history. While the influence of race was highly variable, this effect was found to be statistically significant but statistically small. By contrast, the above effect size analysis indicates that the influence of race in Federal courts was

not statistically significant at conventional levels of significance (i.e., 0.05) and was statistically small.

The finding that the two effect size distributions displayed a statistically greater level of variability than that expected by only sampling error suggests that moderator variables may exist which explain the observed variation in the magnitude of effect sizes. The following analyses investigate how the effect of race varies by pertinent methodological, sample and contextual features of each sentencing context. In the following moderator analyses the dependent variable is the overall effect size (as described above) and all models are analyzed using random (mixed) effects models.

BIVARIATE ANALYSES

Table 13 presents bivariate analyses assessing the relationship between key methodological variables and magnitude of unwarranted racial disparity for both analyses of State and Federal data. The first column under each heading lists the methodological variables analyzed. The second column lists the mean odds ratio, while the third and fourth columns display the lower and upper bounds of the 95 percent confidence interval for each category of the variables examined. The fifth column lists the frequency of each level (category) of each methodological variable. Statistically significant differences between the levels of a variable are indicated by a series of crosses next to the name of the variable (listed in the first column); that is, variables that statistically moderate effect size are denoted by crosses next to the name of that variable. For example, a series of crosses (†††) are listed next to the variable representing the precision of criminal history measure ("Criminal History Level of Measure"); the three crosses indicate that this variable's association with effect size was statistically significant at less than the 0.01 level, whereas the "s" following the crosses denotes that this finding of a statistically

significant relationship is found in analyses of State data (if the relationship had been significant in analyses of Federal data, a "f" would have appeared after the series of crosses). Variables without a series of crosses listed next the variable label (in column one) indicates that this variable do not have a statistically significant relationship to effect size. Statistically significant mean odds ratio effect sizes (i.e., effect sizes statistically greater or less than 1) are indicated by a series of asterisks next to the mean odds ratio (listed in the third column). For instance, the asterisks listed in the first row (beneath the header) denotes that the mean odds ratio (1.64) in analyses of State data that utilized a single dichotomy as a measure of criminal history is statistically different from 1.00 at a level of significance at less than the 0.01 level.

As hypothesized, Table 13 reveals analyses that less precise measures of criminal history and offense seriousness produced larger estimates of unwarranted racial disparity than analyses that used more precise measurements. Specifically, Table 13 indicates that analyses that measured criminal history with only a single dichotomy produced larger effect size estimates than those analyses that used more precise measures (i.e., multiple dichotomies or variables measured at the ordinal level or higher). Likewise, analyses that employed common law offense types as measures of offense seriousness produced larger effect sizes than those analyses that utilized more specific measures of offense seriousness. Both of these observed differences are statistically significant in analyses of State data; analyses of Federal data follow the same pattern, however, the small number of effect sizes reduced the statistical power of the test of differences.

Table 13. African-American/White Contrasts by Methodological Features

	State Data				Federal Data			
		95% C.I.				95% C.I.		
Methodological Feature	Mean Odds Ratio	Lower	Upper	*k*[a]	Mean Odds Ratio	Lower	Upper	*k*
Criminal History Level of Measure†††s								
Single Dichotomy	1.64***	1.43	1.88	16	1.23	0.86	1.76	5
Multiple Dichotomies/Ordinal or Higher	1.21***	1.13	1.29	85	1.11	0.89	1.39	10
Type of Criminal History Measure†††f								
Arrest/Charges Filed	1.21	0.86	1.69	3	1.03	0.63	1.69	1
Conviction	1.18**	1.05	1.32	33	0.97	0.81	1.15	10
Incarceration	1.27***	1.07	1.50	13	——	——	——	0
Composite/Multiple Measures	1.32***	1.21	1.45	48	1.58***	1.22	2.04	3
Type of Offense Seriousness Measure††s								
Common Law	1.43***	1.28	1.59	33	1.30	0.81	2.07	2
Offense Rating	1.22***	1.08	1.37	32	0.95	0.67	1.35	4
Common Law and Offense Rating	1.20***	1.08	1.33	36	1.19	0.97	1.48	9
Controls for Type of Counsel††s								
Yes	1.16**	1.03	1.30	35	1.20	0.75	1.93	2
No	1.33***	1.23	1.44	66	1.14	0.92	1.39	13
Controls for Method of Disposition†s								
Yes	1.20***	1.10	1.31	53	1.12	0.93	3.77	1
No	1.36***	1.24	1.49	48	1.76	0.83	1.35	14
Controls for Selection Bias†††f								
Yes	1.09	0.88	1.35	7	1.65***	1.23	2.21	3
No	1.30***	1.21	1.39	94	1.01	0.89	1.20	12
Controls for Pre-trial Release Status								
Yes	1.21***	1.06	1.38	29	1.23	0.91	1.66	6
No	1.30***	1.21	1.39	72	1.10	0.86	1.39	9
Controls for Victim Injury								
Yes	1.13	0.92	1.39	10	1.56*	0.99	2.45	2
No	1.29***	1.21	1.39	91	1.08	0.89	1.31	13

Table 13. African-American/White Contrasts by Methodological Features (cont.)

	State Data				Federal Data			
		95% C.I.				95% C.I.		
Methodological Feature	Mean Odds Ratio	Lower	Upper	k[a]	Mean Odds Ratio	Lower	Upper	k
Controls for Weapon Possession/Use††s								
Yes	1.14**	1.02	1.28	31	1.11	0.84	1.46	6
No	1.34***	1.24	1.45	70	1.18	0.91	1.52	9
Controls for Defendant SES†††s, †f								
Yes	1.15***	1.03	1.29	35	1.08	0.90	1.29	2
No	1.34**	1.24	1.45	66	1.80**	1.11	2.94	13
Questionable Analysis†s								
Yes	1.17**	1.02	1.33	28	0.95	0.67	1.34	4
No	1.31***	1.22	1.41	72	1.21**	1.00	1.46	11
African-Americans Only†s, †††f								
Yes	1.23***	1.14	1.33	73	1.49***	1.28	1.74	9
No	1.39***	1.24	1.56	28	0.88	0.75	1.03	6
Unpublished††s, †f								
Yes	1.14**	1.03	1.27	36	1.59**	1.07	2.36	2
No	1.35***	1.25	1.46	65	1.07	0.89	1.28	13

a. k = number of effect sizes with non-missing values.

* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between African-Americans and whites.

† $p < .10$, †† $p < .05$, ††† $p < .01$ for test of differences between mean odds ratios; i.e., rejects hypothesis of equality of mean odds ratios between levels of moderator variable.

Note: The "s" following a series of crosses indicates that this relationship was statistically significant at the specified level of significance in analyses of State data. Likewise, the "f" following a series of crosses indicates that this relationship was statistically significant at the specified level of significance in analyses of Federal data

The results of Federal and State data diverge sharply. In regards to sentencing contexts that analyzed State data, generally, the moderator variables describing the various control variables employed in the coded analyses behaved as predicted; in that, with only a few exceptions, the inclusion of these control variables reduced the magnitude of unwarranted racial disparity. Most of these reductions in the size of unwarranted racial disparity were statistically small but several of these differences were statistically significant. Specifically, analyses that controlled for the type of defense counsel, method of disposition, defendant SES, and use/possession of weapons all yielded statistically smaller effect sizes than analyses that omitted these important variables. Interestingly, analyses that employed questionable analytic strategies had a smaller mean effect size than analyses using more appropriate analytic techniques. Further, there is evidence of publication bias in this area of research in that unpublished studies exhibited substantively and statistically smaller effect sizes than published studies.

A parallel analysis was performed using the Federal data (see right side of Table 13). The small number of effect sizes included in this analysis limits the statistical power of this analysis; in spite of this limitation, several noteworthy relationships were revealed. In concordance to the State analyses, the inclusion of controls for defendant SES was found to be negatively associated with unwarranted disparity. Furthermore, as expected from Klepper et al's (1983) arguments, those analyses that employed techniques designed to correct for selection bias produced larger estimates of unwarranted disparity than other types of analyses. Also, as expected, analyses measuring race precisely (i.e., separating African-Americans from other minorities) yielded statistically larger estimates of unwarranted disparity than those that measured race imprecisely. Table 13 also indicates that unpublished analyses and analyses that

employed multiple measures or composite measures of defendant criminal history were associated with larger estimates of unwarranted disparity.

The above analyses indicate that several of the coded methodological features were meaningfully related to magnitude of unwarranted racial disparity at the bivariate level. As a further step toward addressing the relationship between methodological features and size of unwarranted racial disparity, Table 14 reports the results of four bivariate regressions (i.e., correlations) between the number of specific types of variables included in each analysis and effect size. In these bivariate (simple) regressions, the logged odds ratio overall effect size was regressed on each of the continuous moderator variables separately. It is expected that as the number of variables increases, especially the number of control variables, the magnitude of unwarranted racial disparity will decrease.

Table 14a. African-American/White Contrasts by Methodological Variables (Bivariate Regressions)

	State Data				
		95% C.I.			
Number of:	Slope	Lower	Upper	R^2	k^a
Total Variables	-0.004	-0.011	0.002	0.017	101
Criminal History Variables	-0.039	-0.097	0.020	0.029	100
Offense Variables	-0.045	-0.105	0.016	0.042	100
Control Variables***	-0.022	-0.036	-0.007	0.086	100

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

From Table 14 it is apparent that only the number of control variables included in each analysis had a statistically significant negative relationship to size of unwarranted disparity. What's more, this relationship was evident in both analyses of Federal and State court data. In fact, this relationship accounts for a sizeable proportion of the variation in effect size; roughly 9% in State sentencing contexts and nearly 26% in Federal sentencing contexts. These bivariate analyses also reveal that the magnitude of unwarranted racial disparity was not statistically

related to the total number of variables or the number of variables measuring criminal history in either analyses of Federal or State data.

Table 14b. African-American/White Contrasts by Methodological Variables (Bivariate Regressions)

	Federal Data				
		95% C.I.			
Number of:	Slope	Lower	Upper	R^2	*k*
Total Variables	0.002	-0.032	0.035	0.001	15
Criminal History Variables	0.066	-0.021	0.152	0.127	15
Offense Variables*	0.055	-0.003	0.113	0.181	15
Control Variables**	-0.061	-0.115	-0.008	0.257	15

a. *k* = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

The findings from State sentencing indicate that estimates of unwarranted disparity disadvantaging African-Americans were greatest in those analyses utilizing less precise measures and fewer control variables. In regards to analyses of Federal sentencing contexts, the association between methodological rigor and effect size was not nearly as clear. In some regards, precise measures of important variables were negatively related to effect size as expected (e.g., precise measures of race); similarly, the presence of important control variables (e.g., defendant SES) also was negatively related to effect size. Other precise measures of important variables (e.g., criminal history or offense seriousness), however, had no meaningful relationship to effect size in Federal analyses.

The association of methodological variables with effect size makes it important to consider whether the general finding of unwarranted racial disparity disadvantaging African-Americans is completely attributable to methodologically flawed studies. In order to address this question, we imposed a series of increasingly restrictive constraints on the analysis of State level effect sizes; i.e., only findings from State analyses are included. First, we considered the mean

effect size for only those studies that measured criminal history using more precise measures *and* also measured current offense seriousness using more precise measures (i.e., used offense severity ratings or offense severity ratings and common law offense types). Second, we kept these constraints while adding the additional constraint that all analyses had to control for defendant SES. Lastly, we added to the preceding constraints an additional constraint mandating that all analyses had to measure race of defendant precisely (i.e., must separate African-Americans from other racial/ethnic minorities).

These analyses are presented in Table 15. From these analyses, it is apparent that the mean effect size decreased as more restrictions are imposed, but the influence of race remains even after all of these factors have been taken into consideration. Specifically, when the analysis is confined to only those analyses that utilized precise measures of criminal history and offense seriousness, the random effects mean odds ratio decreased only slight (from 1.28 to 1.21). Interestingly, the mean effect size was more strongly influenced by the inclusion of variables measuring defendant SES than by the preceding restrictions; in fact, after restricting the analysis to only those studies that included a measure of defendant SES the mean odds ratio drops to 1.11. After the final restriction was added to the preceding restrictions, only 22 of the 101 State analyses remained; yet even in this reduced sample of studies, the influence of race remained statistically small but statistically significant. Moreover, even in the most constrained model the influence of race continued to vary beyond that expected by chance (sampling error) alone ($Q[21] = 40$, $p = 0.008$). Thus, the influence of race is reduced but remained statistically significant even in the most rigorous analyses, and this influence varied widely.

Table 15. Mean Odds Ratio by Increasingly Methodologically Restrictions

		95% C.I.		
Restrictions	Mean Odds Ratio	Lower	Upper	k^a
Precise Measures of Criminal History & Offense Seriousness	1.21***	1.13	1.29	66
Preceding Plus Measure Defendant SES	1.11***	1.04	1.19	25
Preceding Plus Measure Race Precisely	1.14***	1.06	1.22	22

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

The relationship between effect size and characteristics of each sample was also examined (see Table 16). In the analysis of State data, contrary to our third hypothesis, none of the coded sample characteristics (mean age, proportion female, proportion African-American) were statistically related to magnitude of unwarranted racial disparity. This analysis, however, was somewhat hindered by a significant amount of missing data; that is, many analyses did not report basic descriptive statistics describing the sample under examination. This lack of information is most apparent in regards to mean age of sample, where only 41 of the 101 sentencing contexts reported mean age of the sample.

Table 16a. African-American/White Contrasts: Log Odds Ratio by Sample Characteristics (Bivariate Regressions)

	State Data				
		95% C.I.			
Sample Characteristic	Slope	Lower	Upper	R^2	k^a
Mean Age	-0.008	-0.034	0.017	0.009	41
Proportion Female***	-0.129	-0.541	0.282	0.004	78
Proportion African-American	-9.4e-05	-0.003	0.003	0.000	96

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

In the bivariate analysis of the Federal sentencing studies, as the proportion of females in a sample increases the magnitude of unwarranted racial disparity decreases substantially. This finding suggests that unwarranted racial disparity is most pronounced in samples focusing on

sentencing patterns among males. This finding comports with recent findings from primary research (Steffensmeier et al. 1998; Spohn and Holleran 2000).

Table 16b. African-American/White Contrasts: Log Odds Ratio by Sample Characteristics (Bivariate Regressions)

	Federal Data				
		95% C.I.			
Sample Characteristic	Slope	Lower	Upper	R^2	*k*
Mean Age	-0.037	-0.202	0.127	0.030	7
Proportion Female	-0.623	-1.045	-0.200	0.365	12
Proportion African-American	0.546	-0.550	1.641	0.063	15

a. *k* = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Table 17 investigates the relationship between the coded contextual variables and effect size. Analyses of State data indicate that while analyses conducted with city/county level data produced noticeably larger effect sizes than analyses conducted with state level data (i.e., data pooled from all jurisdictions within a state), this difference was not statistically significant when the "other" category was included. However, removing these miscellaneous "other" sentencing contexts from the analysis shows this difference was statistically significant ($p = 0.09$); although the effect was small. This finding suggests that analyses which pool all state level data into one data set may suffer from aggregation bias. That is, more (or less) unwarranted disparity may be apparent in disaggregated data than in pooled data sets (see Nelson, 1992, 1995 for an example of this phenomenon).

Unwarranted racial disparity disadvantaging African-Americans was larger in Southern jurisdictions than in non-Southern jurisdictions; however, this difference also was not statistically significant at conventional levels of statistical significance ($p = 0.13$). Perhaps most interestingly, at the bivariate level, in analyses of State sentencing contexts, jurisdictions employing structured sentencing displayed smaller unwarranted racial disparity disadvantaging

African-Americans. This difference, however, is statistically small and falls short of conventional statistical significance ($p = 0.06$).

Table 17a. African-American/White Contrasts by Contextual Characteristics

	State Data			
		95% C.I.		
Contextual Feature	Mean Odds Ratio	Lower	Upper	k[a]
Jurisdiction Type				
City/County	1.35***	1.23	1.48	51
State	1.20***	1.10	1.32	47
Other	1.36	0.93	1.99	3
Structured Sentencing[†]				
Yes	1.18***	1.06	1.31	32
No	1.34***	1.23	1.45	69
Before 1980				
Yes	1.28***	1.15	1.43	44
No	1.28***	1.18	1.39	56
After Commencement of Drug War				
Yes	1.24***	1.09	1.42	23
No	1.29***	1.20	1.40	77
Southern				
Yes	1.41***	1.22	1.64	25
No	1.25***	1.15	1.34	75

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between African-Americans and whites.
† $p < .10$, †† $p < .05$, ††† $p < .01$ for test of differences between mean odds ratios; i.e., rejects hypothesis of equality of mean odds ratios between levels of moderator variable.

Table 17a also shows that time period of data collection was not substantively or statistically related to effect size in State analyses. Specifically, the mean odds ratio effect size was nearly identical in analyses that collected data before and after 1980. Furthermore, analyses utilizing data collected after 1986 (i.e., during the drug war) did not display greater signs of unwarranted racial disparity than those conducted before 1986. In addition to these dichotomous measures of time period, we also investigated the association between effect size and time period by correlating the midpoint of each data series with effect size using the meta-analytic analog to

simple (bivariate) regression. This analysis (not shown) continued to indicate that time period was not statistically related to effect size; the unstandardized regression coefficient between the two measures was 0.003 ($p = 0.47$).

The parallel analysis of Federal sentencing contexts diverges sharply from the results of State level sentencing contexts. Specifically, more recent analyses (i.e., those analyses conducted using data after the implementation of the Federal sentencing guidelines) found much stronger evidence of unwarranted racial disparity than analyses from earlier time periods. In particular, analyses of Federal sentencing practices conducted with data collection since 1980 had a mean effect size of 1.58 in comparison to a mean effect size of 1.02 from earlier studies. Moreover, all three contextual measures pertinent to Federal analyses are completely confounded; because all three independent analyses conducted with Federal data collected after 1980 are the same three analyses conducted since the commencement of the drug war and these three analyses are also the only three independent analyses of Federal data conducted since the implementation of the Federal guidelines. Thus, the independent effects of these contextual variables in the present data set are inseparably intertwined, and therefore no conclusions regarding these variables can be made—other than stating that analyses of more recent Federal data yield considerably stronger evidence of unwarranted racial disparity than earlier analyses of Federal data.

Table 17b. African-American/White Contrasts by Contextual Characteristics

	Federal Data			
		95% C.I.		
Contextual Feature	Mean Odds Ratio	Lower	Upper	*k*
Structured Sentencing††				
Yes	1.58	1.18	2.11	3
No	1.02	0.86	1.22	12
Before 1980††				
Yes	1.02	0.86	1.22	12
No	1.58	1.18	2.11	3
After Commencement of Drug War††				
Yes	1.58	1.18	2.11	3
No	1.02	0.86	1.22	12

a. *k* = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between African-Americans and whites.
† $p < .10$, †† $p < .05$, ††† $p < .01$ for test of differences between mean odds ratios; i.e., rejects hypothesis of equality of mean odds ratios between levels of moderator variable.

The findings from analyses of State level data yield support for our fourth hypothesis. In that, jurisdictions with structured sentencing displayed smaller amounts of unwarranted racial disparity than those sentencing contexts without such mechanisms. By contrast, the analyses of Federal sentencing contexts indicate that structured sentencing was associated with larger unwarranted racial disparity. This finding, however, is completely confounded with other variables which may be responsible for this association; i.e., this association may be spurious.

The above analyses all examined unwarranted racial disparity in sentencing outcomes utilizing the *overall* measure of unwarranted racial disparity, which combines offense and outcome specific measures of unwarranted disparity into one global measure of unwarranted disparity. The next set of analyses examines variations in effect size by specific types of offenses and by specific types of sentencing outcomes. It is important to recognize that while the following effect sizes are specific in one regard, either specific to offense or type of sentencing outcome, they are not specific in both regards. That is, the offense specific analyses combine

effect sizes from the five different types of sentencing outcomes (by taking the weighted average, when a sentencing context reported analyses of multiple types of sentencing outcomes). Similarly, the sentencing outcome specific analyses combine effect sizes from the various offense types.

Table 18a shows that 19 State sentencing contexts conducted analyses specific to drug offenses, 21 State sentencing contexts conducted analyses specific to property offenses, and 36 State sentencing contexts conducted analyses specific to violent offenses. Comparing the mean odds ratio effect sizes from these offense specific analyses indicate that unwarranted racial disparity was greatest in regards to drug offenses, as hypothesized. Specifically, the overall random effects mean effect size for drug offenses was 1.40, which is noticeably greater than the mean effect size for either property offenses (1.09) or violent offenses (1.20).

Table 18a. African-American/White Contrasts by Type of Offense

	State Data			
		95% C.I.		
Offense Type	Mean Odds Ratio	Lower	Upper	k^a
Drug Offense	1.40***	1.21	1.62	19
Property Offense	1.09	0.95	1.25	21
Violent Offense	1.20***	1.07	1.34	36

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between African-Americans and whites.

Analyses of Federal data exhibited a different pattern of results. Unwarranted racial disparity was greater in analyses of property offenses than drug offenses. Somewhat surprisingly, the magnitude of unwarranted racial disparity in Federal drug offenses was not statistically significant

Table 18b. African-American/White Contrasts by Type of Offense

	Federal Data			
		95% C.I.		
Offense Type	Mean Odds Ratio	Lower	Upper	*k*
Drug Offense	1.08	0.85	1.38	7
Property Offense	1.32***	1.15	1.52	7

a. *k* = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between African-Americans and whites.

Note that no statistical significance tests were performed on these comparisons of offense specific analyses, as these effect sizes are not statistically independent. Yet, at least in State sentencing contexts, there appears to be substantive differences in the amount of unwarranted racial disparity by type of offense.

Table 19 examines variation in effect size by type of sentencing outcome. This analysis was confined to only the 101 analyses of State sentencing contexts, as the number of Federal sentence contexts (15) is too small to support separating these effect sizes into five categories. The most common type of sentencing outcome involved imprisonment decisions ($k = 64$), followed by decisions regarding length of incarcerative sentence; a considerably smaller number of effect sizes examined other types of sentencing decisions. From Table 19 it is apparent that the random effects mean odds ratio effect size is statistically greater than one for three types of sentencing outcomes: imprisonment decisions, incarcerative sentence length decisions, and decisions relating to discretionary lenience—meaning that unwarranted racial disparity disadvantaging African-Americans was statistically greater than chance for these outcomes. Likewise, the mean effect size from analyses of discretionary harshness outcomes appears to be substantively significant, but the small number of effect sizes reduces the statistical power of this analysis. By contrast, the mean effect size for simultaneous imprisonment/sentence length decisions was neither substantively, nor statistically significant.

Perhaps more importantly, Table 19 also indicates that unwarranted racial disparity disfavoring African-Americans is greatest in regards to imprisonment decisions. Specifically, the mean odds ratio in analyses of imprisonment decisions was 1.34, whereas the mean odds ratio was 1.17 in analyses of sentence length decisions. Substantially fewer analyses examined other types of sentencing outcomes. Among these other sentencing outcomes, the type of sentencing outcome with the smallest estimate of unwarranted racial disparity were simultaneous analyses of both imprisonment and sentence length outcomes (1.10). Moderate estimates of unwarranted disparity were evident in examinations of discretionary decisions.

Table 19. African-American/White Contrasts by Outcome Type

		95% Confidence Interval		
Outcome Type	Mean Odds Ratio	Lower	Upper	*k*[a]
Outcome Specific Effect Sizes				
Imprisonment	1.34***	1.24	1.45	64
Length of Incarcerative Sentence	1.17***	1.09	1.27	50
Simultaneous Imprisonment/Length	1.10	0.96	1.27	15
Discretionary Lenience	1.24***	1.07	1.45	12
Discretionary Harshness	1.21	0.83	1.76	6
Overall Effect Sizes				
Imprisonment	1.38***	1.25	1.54	34
Length of Incarcerative Sentence	1.23***	1.07	1.42	19
Simultaneous Imprisonment/Length	1.05	0.85	1.31	11
Discretionary Harshness	1.73***	1.23	2.42	3
Mixture of the Above Types	1.22***	1.10	1.35	34

a. *k* = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between African-Americans and whites.

Another way to assess whether type of sentencing outcome is related to effect size is to categorize each of the overall effect size measures into one of six types of effect sizes; effects sizes that concerned: 1) only imprisonment decisions, 2) only sentence length decisions, 3) only simultaneous imprisonment/length decisions, 4) only discretionary leniency, 5) only

discretionary harshness, and 6) mixed two or more of the preceding types of sentencing outcomes. The advantage of this model is that statistical significance testing can be conducted on these effect sizes, as they are independent. The disadvantage of this model is that few analyses examined unwarranted racial disparity in regards to only discretionary outcomes (i.e., discretionary lenience or harshness).

The bottom section of Table 19 presents the results of this alternative model. This analysis also indicates that type of sentencing outcome is associated with effect size. Further, this alternative model continues to indicate that unwarranted racial disparity was greatest for imprisonment and discretionary decisions, and smallest in analyses that examined length of incarcerative sentences or simultaneously examined imprisonment and sentence length. Global significance testing indicates that effect size is not equivalent among these sentencing outcome types. Pairwise contrasts indicated that the mean effect size for imprisonment and discretionary harshness sentencing outcomes were statistically different from analyses using simultaneous imprisonment/sentence length outcomes, and mean effect sizes from imprisonment and discretionary sentencing decisions were marginally different for analyses with mixed sentencing outcome types.

The above analyses indicate that type of sentencing outcome is another important source of variation in unwarranted racial disparity in sentencing. These analyses suggest that unwarranted racial disparity was greatest in imprisonment decisions and in decisions related to discretionary leniency. By contrast, unwarranted racial disparity was smallest in analyses that simultaneously assessed imprisonment and sentence length decisions.

we also attempted to conduct the preceding bivariate analyses separately for each sentencing outcome. The limited number of outcome specific effect sizes, constrained these

analyses to only imprisonment and sentence length decisions. Substantively, bivariate analyses of imprisonment decisions were nearly identical to the preceding analyses of the overall effect size. The bivariate analyses of sentence length effect sizes substantively were similar to those presented above; however, few bivariate relationships were not statistically significant. Exceptions to this general pattern of similarity were that analyses utilizing questionable analytic techniques and analyses using more precise measures of race produced statistically larger effect sizes than other analyses. Another dissimilarity was that when only sentence length decisions were considered, analyses conducted using data collected before the sentencing reform movement (i.e., before 1980) produced statistically larger effect sizes than analyses examining data collected after 1980.

MULTIVARIATE MODEL

So far several variables have been identified that have exhibited bivariate relationships with effect size. The following analysis estimated a multivariate model that attempted to find the factors associated with variation in effect size. This multivariate model was restricted to only those effect sizes calculated from analyses of State data, as there are too few effect sizes from Federal sentencing contexts to support such an analysis.

This multivariate analysis began by entering all of the moderator variables that exhibited a statistically significant relationship to effect size, including type of sentencing outcome (representing by a series of dummy variables), into an initial multivariate model. Model 1 in Table 20 presents the results of this analysis. While the model is statistically significant and accounts for a large portion of the variation effect size (36%), few of the variables are statistically significant. In fact, only two variables were statistically significant; precision of criminal history measure (i.e., analyses employing multiple dichotomous measures of criminal

history or criminal history measures scaled at the ordinal level or higher are coded as "1") and type of sentencing outcome analyzed (the reference category is sentencing outcomes related to imprisonment decisions). This model, however, runs a substantial risk of overfitting the data, as the ratio of observations (i.e., effect sizes) to moderator variables is rather low (5.6). There is also some evidence of multicollinearity in this model as the bivariate correlation between the number of control variables included in each analysis and the presence of control variables for defendant SES was rather high (r = 0.61). Additional evidence of multicollinearity is found by removing either one of these variables from the model, in that the effect of doing so is to increase the other variable's relationship with the dependent variable.

Table 20. African-American/White Contrasts: Multivariate Regression

	Model 1		Model 2	
Variable	b	*p*	b	*p*
Number of Control Variables	-0.004	0.648	——	——
Precise Criminal History Measure	-0.230	0.032	-0.357	0.000
Offense Rating (Offense Seriousness Measure)	-0.124	0.222	——	——
Common Law & Offense Rating (Offense Seriousness Measure)	-0.158	0.126	——	——
Mode of Disposition	0.188	0.066	——	——
Type of Defense Counsel	-0.163	0.089	——	——
Possession/Use of Weapons	-0.010	0.900	——	——
Defendant SES	-0.023	0.768	-0.119	0.060
Questionable Analysis	0.001	0.918	——	——
African-Americans Only (Race Measure)	-0.006	0.936	——	——
Unpublished Analysis	-0.051	0.460	-0.101	0.080
Structured Sentencing	-0.118	0.142	——	——
State Level Data	-0.125	0.151	——	——
Sentence Length (Outcome Type)	0.046	0.630	0.041	0.640
Simultaneous Imprisonment/Length (Outcome Type)	-0.118	0.440	-0.062	0.603
Discretionary Harshness (Outcome Type)	0.668	0.001	0.404	0.006
Mixed Sentencing Outcomes (Outcome Type)	0.153	0.069	0.107	0.149
Constant	0.544	0.000	0.539	0.000
k[a]	96		96	
R^2	0.36***		0.30***	

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Model 2 in Table 20 is a post-hoc model comprised of variables accounting for the maximum amount of variation in the overall effect size measure of unwarranted racial disparity. This model indicates that the strongest predictor of effect size was precision of criminal history measure; once again, studies that employed either multiple measures of criminal history or measures scaled at the ordinal level or higher were associated with substantially smaller effect sizes than analyses utilizing only a single simple dichotomy as a measure of criminal history. The multivariate model continues to indicate that the presence of controls for defendant SES was negatively associated with effect size. Additional models were estimated that included the variable measuring number of control variables included in each primary analysis, instead of the moderator variable flagging primary analyses that controlled for defendant SES; however, this substitution led to a marginal reduction in R^2 (0.28 vs. 0.30). Furthermore, including both variables in the same mode results in neither variable attaining statistical significance, and the inclusion of both variables produces a negligible increase the proportion of variation accounted for by this model (0.307 vs. 0.301).

This model also reveals type of sentencing outcome had a substantial relationship to effect size. After controlling for methodological differences between analyses, sentencing outcomes that examined imprisonment decisions, sentence length, simultaneous measures of imprisonment and sentence length, or a mixture of the preceding all produced comparable effect sizes; however, those analyses that analyzed discretionary punitiveness were associated with considerably larger effect sizes than other types of sentencing outcomes. Lastly, even after controlling for other important factors, unpublished studies produced somewhat smaller estimates of unwarranted racial disparity than published studies.

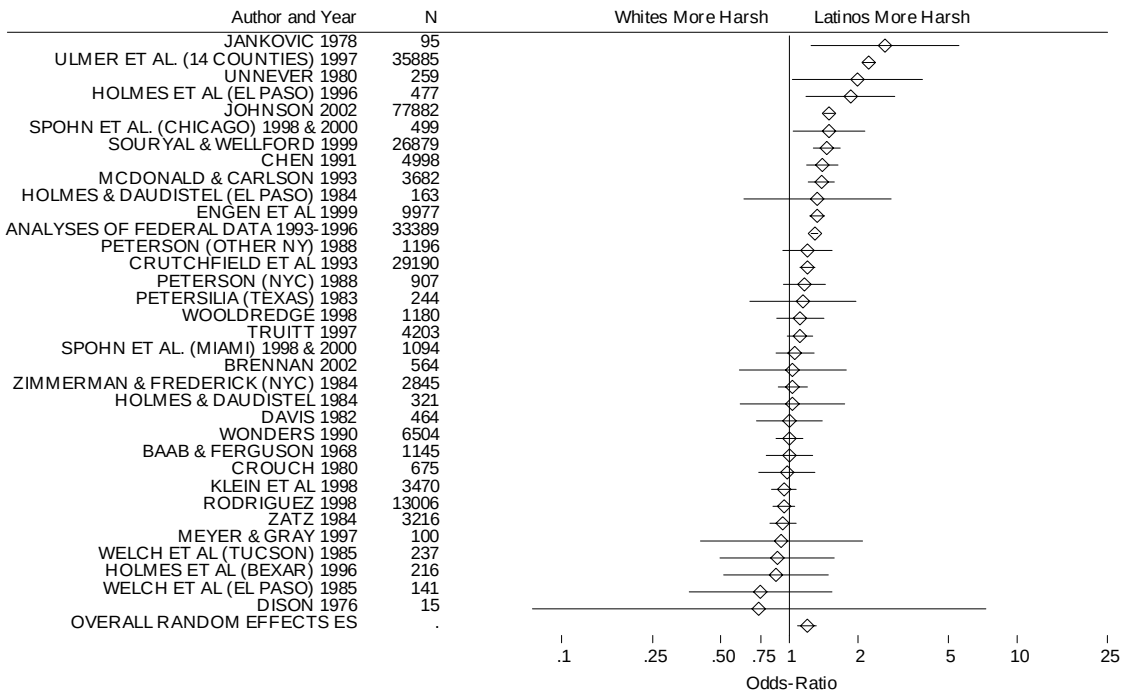
Latino/White Contrasts

Thirty-four of the 122 sentencing contexts compared sentencing outcomes of Latinos to those of whites. These 34 sentencing contexts estimated 109 separate Latino/white contrasts. Seventy-two percent of these effect sizes were greater than 1, indicating that Latinos were sentenced more harshly than whites independent of criminal history and offense seriousness. Moreover, 48% of the 109 effect sizes were statistically greater than 1. By contrast, 1% of Latino/white effect sizes were statistically less than 1, indicating that whites were significantly disadvantaged at sentencing in comparison to Latinos. These findings strongly suggest that defendant ethnicity matters in sentencing; once again, however, drawing firm conclusions is premature given that these analyses are not statistically independent.

In the same manner as before, we calculated the *overall* effect size for Latino/white sentencing contrasts. Figure 4 displays each of these 34 effect sizes in a forest plot. From this plot, it is apparent that the effect sizes showed considerable variability, ranging from a low of 0.74 (a small negative relationship) to a high of 2.64 (a modestly large positive relationship). Furthermore, the *Q* statistic confirms that this sample of effect sizes exhibits a statistically greater amount of variability than would be expected by sampling error alone ($Q[33] = 471$, $p < 0.0001$). Interestingly, 35% of the 34 independent overall effect sizes were statistically greater than 1, yet none of the 34 effect sizes were statistically less than 1—indicating that in none of the sentencing contexts were whites sentenced significantly more harshly than Latinos. The bottom of Figure 4 graphically displays the random effects mean random effects odds ratio. From this figure it is apparent that the overall mean odds ratio is statically significant, as the confidence interval for this mean effect size does not cross the line representing an odds ratio of one. Note that two of the 34 sentencing contexts assessed sentencing practices in the Federal system.

These two effect sizes are dropped from *all* subsequent analyses, in order to guard against systematic differences between Federal and State jurisdictions distorting the cumulative findings of this research.

Figure 4. Forest Plot of Latino/White Contrasts



The overall mean odds ratio for Latino/white contrasts for the 32 State effect sizes is displayed in Table 21. This table indicates that the estimated mean odds ratio was 1.18, which is statistically significant. This mean odds ratio, assuming a 50% rate of punishment for whites translates into a punishment rate of approximately 54% for Latinos. Thus, while this mean effect size is statistically significant, the magnitude of this mean effect size is statistically small; this effect, however, varies from context to context. Furthermore, the distribution of Latino/white

contrasts does not appear to contain any extreme outliers (see Figure 5 – logged odds ratios are displayed); therefore, this finding of a statistically small ethnic effect is not attributable to the distorting effect of an outlier.

Table 21. Latino/White Contrasts by Type of Effect Size Analysis

	Mean	95% C.I.			
Type of Analysis	Odds Ratio	Lower	Upper	*Q*	*k*[a]
Fixed Effects	1.33***	1.32	1.39	467***	32
Random Effects	1.18***	1.05	1.33		32
Unweighted Fixed Effects	1.17	0.83	1.66	0	32

a. *k* = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between Latinos and whites.

Figure 5. Latino/White Contrasts: Stem and Leaf Plot

-3*	1
-2*	9
-1*	32
-0*	87662
0*	0003336
1*	114588
2*	88
3*	38
4*	00
5*	
6*	29
7*	
8*	0
9*	7

BIVARIATE ANALYSES

The variability exhibited in these effect sizes suggests that there may be systematic variation in effect size, the analyses that follow attempt to account for the observed variation in results using these coded moderator variables. The moderator variable analysis follows much the same format as that utilized in the analysis of African-American effect sizes. We begin by

assessing the bivariate relationship between the methodological moderator variables and the overall effect size measure of unwarranted sentencing disparity. Table 22 displays the results of these analyses. These results are similar in many regards to those from the analysis of African-American/white contrasts; however, the smaller number of effect sizes reduces the statistical power of this analysis, and as a result few of the observed differences are statistically significant. For example, once again analyses that employed less precise measures of criminal history and seriousness of present offense yielded larger estimates of unwarranted sentencing disparity than analyses that used more precise measures. Specifically, analyses that controlled for criminal history using only a single dichotomous variable produced larger estimates of unwarranted disparity than analyses that employed specific measures; this difference, however, is not statistically significant. Likewise, as hypothesized, analyses that controlled for type of defense counsel, pre-trial release status, victim injury, and possession/use of a weapon all produced substantively smaller effect sizes than analyses that omitted these control variables. None of these differences are statistically significant at conventional levels of significance; thus, support for hypothesis two in these data is weak but the direction and magnitude of differences between methodologically sophisticated and methodologically weak studies does provide some support for hypothesis two.

Table 22. Latino/White Contrasts by Methodological Variables

	Mean	95% Confidence Interval		
Methodological Feature	Odds Ratio	Lower	Upper	k^a
Criminal History Level of Measure				
Single Dichotomy	1.30*	0.96	1.74	6
Multiple Dichotomies/Ordinal or Higher	1.18**	1.07	1.29	26
Type of Criminal History Measure				
Arrest/Charges Filed	1.24	0.93	1.65	2
Conviction	1.30***	1.14	1.48	14
Incarceration	1.17	0.90	1.52	4
Multiple	1.07	0.93	1.23	12

Table 22. (cont.) Latino/White Contrasts by Methodological Variables

Methodological Feature	Mean Odds Ratio	95% Confidence Interval Lower	Upper	k[a]
Type of Offense Seriousness Measure				
Common Law	1.06	0.85	1.33	6
Offense Rating	1.09	0.90	1.31	9
Common Law and Offense Rating	1.25***	1.12	1.39	17
Controls for Type of Counsel				
Yes	1.12*	0.97	1.30	13
No	1.23***	1.10	1.38	19
Controls for Method of Disposition				
Yes	1.22***	1.10	1.36	21
No	1.09	0.93	1.29	11
Controls for Selection Bias†				
Yes	1.37**	1.15	1.64	5
No	1.14**	1.03	1.25	27
Controls for Pre-trial Release Status				
Yes	1.09	0.95	1.25	15
No	1.25***	1.12	1.40	17
Controls for Victim Injury				
Yes	1.09	0.85	1.40	3
No	1.20***	1.09	1.32	29
Controls for Weapon Possession/Use				
Yes	1.09	0.94	1.26	10
No	1.24***	1.12	1.38	22
Controls for Defendant SES				
Yes	1.14*	0.99	1.32	12
No	1.21***	1.09	1.36	20
Questionable Analysis				
Yes	1.07	0.90	1.27	9
No	1.25***	1.14	1.37	22
Unpublished				
Yes	1.18*	0.99	1.32	11
No	1.22***	1.09	1.36	21

a. k = number of effect sizes with non-missing values.

* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between Latinos and whites.

† $p < .10$, †† $p < .05$, ††† $p < .01$ for test of differences between mean odds ratios; i.e., rejects hypothesis of equality of mean odds ratios between levels of moderator variable.

Bivariate analyses assessing the relationship between the continuous methodological moderator variables and effect size continue to provide only weak support for hypothesis two (see Table 23). From this table, it is apparent that the total number of variables, number of variables measuring offense seriousness, and number of control variables included in each analysis are not statistically related to effect size. The only methodological variable displaying a statistically significant relationship to effect size is the number of variables measuring criminal history. Not only is this negative relationship statistically significant, it is also substantively large—accounting for nearly 12% of variation in effect size.

Table 23. Latino/White Contrasts: Bivariate Regression Methodological Variables

		95% Confidence Interval			
Method Variable	Slope	Lower	Upper	R^2	k^a
Total Variables	-0.010	-0.022	0.002	0.078	32
Criminal History Variables*	-0.036	-0.072	0.001	0.115	32
Offense Variables	-0.006	-0.053	0.041	0.002	32
Control Variables	-0.010	-0.026	0.007	0.043	32

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Continuing with the bivariate analyses, Table 24 assesses the relationship between sample characteristics and effect size. Once again, there is a significant amount of missing data regarding the mean sample age. In spite of this problem, it is apparent that the association between effect size and mean age of sample is negligible. It is also obvious from this analyses that the proportion of female offenders and proportion of Latinos in each sample also were not associated with effect size in any meaningful manner. Thus, hypothesis three clearly is not supported in these analyses.

Table 24. Latino/White Contrasts: Bivariate Regression Sample Characteristics

		95% Confidence Interval			
Sample Characteristic	Slope	Lower	Upper	R^2	k^a
Mean Age	0.004	-0.016	0.023	0.007	17
Proportion Female	-1.7e-04	-0.669	0.669	0.000	30
Proportion Latino	-8.5e-04	-0.004	0.003	0.008	32

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

The coded contextual characteristics displayed a greater level of association with effect size than the other moderator variables analyzed thus far. Time period of data collection was the strongest predictor of effect size among the contextual variables (see Table 25). Those sentencing contexts that analyzed data collected in earlier time periods (before 1980 and before the commencement of the War on Drugs) produced smaller estimates of unwarranted ethnic disparity than those sentencing contexts that analyzed data sets from later periods. This finding is somewhat surprising given the apparent progress America has made in regards to race/ethnic relations in the past several decades. It may be the case that the Drug War has fueled sentencing disparities between Latinos and whites. Another interesting finding is that sentencing contexts located in the Southwestern United States (defined as California, Arizona, New Mexico, Nevada, and Texas) displayed less unwarranted disparity disadvantaging Latinos than sentencing contexts analyzing data from other regions. Contrary to hypothesis four, jurisdictions utilizing structured sentencing mechanisms displayed somewhat larger unwarranted ethnic disparities than those without such mechanisms. Also, in contrast to the analysis of African-American/white effect sizes, there was no relationship between effect size and type of jurisdiction analyzed; that is, sentencing contexts analyzing data collected from counties yielded similar results as those sentencing contexts analyzing pooled state level data.

Table 25. Latino/White Contrasts by Contextual Variables

Contextual Characteristic	Mean Odds Ratio	95% Confidence Interval Lower	95% Confidence Interval Upper	k[a]
Jurisdiction Type				
City/County	1.18**	1.03	1.34	18
State	1.20***	1.06	1.35	14
Structured Sentencing				
Yes	1.24***	1.11	1.40	14
No	1.11	0.97	1.27	18
Before 1980				
Yes	1.09	0.94	1.27	14
No	1.23***	1.11	1.37	18
After Commencement of Drug War††				
Yes	1.31***	1.16	1.48	13
No	1.09	0.97	1.22	19
Southwestern				
Yes	1.09	0.96	1.25	17
No	1.26***	1.12	1.41	15

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between Latinos and whites.
† $p < .10$, †† $p < .05$, ††† $p < .01$ for test of differences between mean odds ratios; i.e., rejects hypothesis of equality of mean odds ratios between levels of moderator variable.

Table 26 data shows that effect sizes varied by type of offense. While only a modest proportion of sentencing contexts analyzed the influence of ethnicity by distinct offense types, the available evidence strongly suggests that the influence of ethnicity was largest in drug offenses (mean odds ratio = 2.01). An odds ratio of this magnitude translates in a 17% difference between Latinos and whites, if we assume a 50% rate of punishment for whites; that is, based on this mean effect size 67% of Latinos would be punished in comparison to 50% of whites. By contrast, the property offense specific mean effect size was more modest; if we continue to assume a 50% rate of punishment for whites, then 58% of Latinos were expected to be punished. Thus, once again, it appears that minorities were most disadvantaged in offenses involving drugs and least disadvantaged in regards to property offenses.

Table 26. Latino/White Contrasts by Type of Offense

Offense Type	Mean Odds Ratio[b]	95% Confidence Interval Lower	Upper	k[a]
Drug Offenses	2.01***	1.57	2.58	13
Property Offenses	1.38**	1.10	1.73	15
Violent Offenses	1.50***	1.21	1.87	18

a. k = number of effect sizes with non-missing values.
b. Random effects odds ratio are displayed as the homogeneity assumption was not met.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between Latinos and whites.

In regards to unwarranted ethnic disparity by sentencing outcome, just as in the African-American/white contrasts, the results of the bivariate analyses indicate that unwarranted disparity in sentencing outcomes was greatest in imprisonment and discretionary decisions (see Table 27). By contrast, ethnicity appears to have a small, perhaps negligible influence, on sentencing decisions involving length of incarcerative sentence and in analyses that combined imprisonment and sentence length decisions.

Table 27. Latino/White Contrasts by Outcome Type

Outcome Type	Mean Odds Ratio	95% Confidence Interval Lower	Upper	k[a]
Outcome Specific Effect Sizes				
Imprisonment Decisions	1.37***	1.16	1.60	20
Length of Incarcerative Sentence Decisions	1.09	0.90	1.33	16
Simultaneous Imprisonment/Length Decisions	1.12	0.98	1.27	9
Discretionary Leniency	1.31	0.85	2.02	5
Discretionary Harshness	1.39	0.70	2.78	3
Overall Effect Sizes				
Imprisonment Decisions	1.26**	1.02	1.56	8
Length of Incarcerative Sentence Decisions	0.97	0.75	1.25	4
Simultaneous Imprisonment/Length Decisions	1.11	0.85	1.45	4
Mixture of Above Types	1.23***	1.10	1.37	16

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .0$ for test of mean odds ratio is equal to 1; i.e., rejects hypothesis of equality of sentencing severity between Latinos and whites.

MULTIVARIATE ANALYSIS

Table 28 presents the findings of a multivariate model that attempts to find the factors associated with variation in effect size. This multivariate model was limited to the 32 effect sizes calculated from analyses of State data. The multivariate model regresses the logged odds ratio overall effect size on the three moderator variables (presence of selection bias corrections, number of criminal history variables, and data collected after the commencement of the War on Drugs) that exhibited a substantive relationship to effect size in the bivariate analyses. Table 28 indicates that this post-hoc three variable model was statistically significant and accounts for a sizeable portion of the variation in effect size (28%). However, none of the variables individually were statistically significant at a conventional significance level. The predictor with the largest unstandardized regression coefficient was the indicator of analyses with corrections for selection bias. Analyses utilizing such analytic procedures yielded substantially larger estimates of unwarranted ethnic disparity than those without such procedures. Another variable with a large unstandardized regression coefficient was time period of data collection (before or after commencement of Drug War); as expected, analyses conducted after the start of the War on Drugs yielded larger estimates of unwarranted ethnic disparity. Lastly, the multivariate analysis indicated that as the number of variables measuring defendant prior criminal history included in the primary author's analysis increased the amount of unwarranted disparity decreased. Other multivariate models (not shown) that included more variables, and different sets of variables were not able to meaningfully improve on the model displayed in Table 28.

Table 28. Latino/White Contrasts: Multivariate Regression

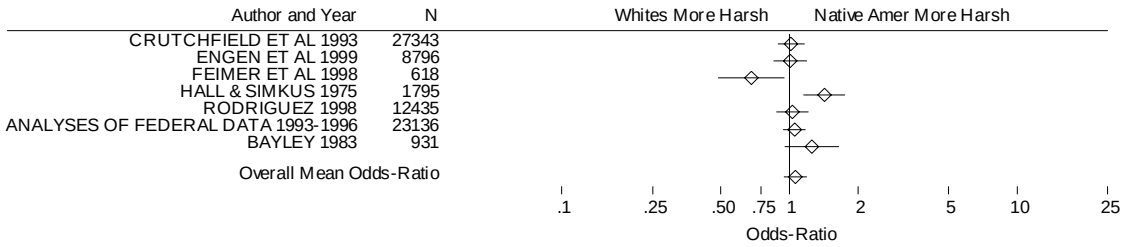
Variable	b	*p*
Selection Bias	0.144	0.114
Number of Criminal History Variables	-0.026	0.111
After Commencement of Drug War	0.140	0.078
Constant	0.133	0.071
k[a]	32	
R^2	0.28***	

a. k = number of effect sizes with non-missing values.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Native-American/White Contrasts

Only seven analyses contrasted sentencing outcomes of Native-Americans to whites (see Figure 6); thus, too few effect sizes were available to support a meta-analytic synthesis of these analyses. Yet, it is interesting to note that six of the seven effect sizes indicated that Native-Americans were sentenced more harshly than whites (86%). Most of these effect sizes were very small; only two effect sizes were statistically significant—one effect size was statistically greater than 1 and one effect size was statistically less than 1. Given this distribution of effect sizes it is not surprising that the random effects mean odds ratio was neither substantively, nor statistically significant; the mean odds ratio was 1.06 with a confidence interval spanning from 0.95 to 1.19.

Figure 6. Forest Plot: Native American/White Contrasts



Asian/White Contrasts

Unfortunately, there were also too few Asian/white contrasts to justify a meta-analytic synthesis. The search strategy uncovered only four independent analyses of Asian/white contrasts. The four overall effect size measures of unwarranted sentencing disparity for these contexts are displayed in a forest plot (see Figure 7). As Figure 7 indicates one of the effect sizes was less than 1, indicating that in this sentencing context whites were sentenced more harshly than Asians; this effect size was not statistically significant, however. The other three effect sizes were all positive, but only one was statistically significant. It should be noted that the first

three effect sizes displayed in Figure 7 all were calculated from sentencing practices in the state of Washington; whereas the fourth effect size examined sentencing practices Federal courts.

Figure 7. Forest Plot: Asian/White Contrasts

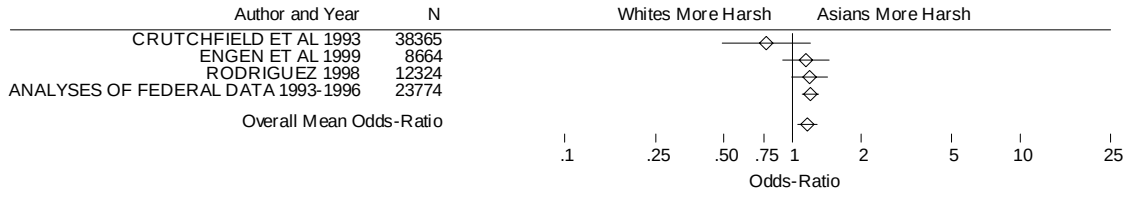


Figure 7 also shows that the mean odds ratio effect size for the overall measure of unwarranted racial disparity. The fixed effects ($Q[3] = 3.80$, $p = 0.288$) mean odds ratio effect size for Asian/white contrasts was 1.18 and the random effects mean odds ratio was 1.16; both mean odds ratios indicate that Asians were at a statistically small disadvantage in sentencing outcomes, after accounting for defendant criminal history and seriousness of present offense. It appears, however, that this finding of statistically significant unwarranted racial disparity disadvantaging Asians is solely attributable to the effect size calculated from Federal court data.

Once this effect size was removed, the mean odds ratio effect size drops to 1.10 and is not statistically significant. Thus, it appears that Asians in Washington State were not sentenced more punitively than whites; yet, Asians in Federal courts were at a statistically significant but statistically small disadvantage in comparison to whites.

Summary of Effect Size Analyses

The preceding effect size analyses revealed that even after taking into consideration current offense seriousness and defendant prior criminal conduct, African-Americans and Latinos were sentenced more harshly, on average, than whites. Specifically, the effect size analyses indicated that sentencing disparity between African-Americans and whites were greater in State sentencing contexts than in Federal sentencing contexts. The mean overall odds ratio for State sentencing contexts was 1.28, whereas the mean overall effect size for Federal sentencing contexts was 1.18. It is very important to note, however, that more recent analyses of Federal data yield considerably greater indications of unwarranted racial disparity (i.e., analyses of Federal data collected since 1980 the mean odds ratio effect size was 1.58 compared to a mean of 1.02 for analyses conducted with data before 1980). Translating these mean odds ratios in percentages while assuming a 50% rate of punishment for whites suggests that African-Americans in State courts would be punished at a rate of 56% and African-Americans in Federal courts would be punished at a rate of 53%. Restricting our attention to only more recent analyses of Federal data would increase this percentage of African-Americans expected to be punished to 61%.

Similarly, the effect size analysis of Latino/white contrasts also reveals that Latinos were sentenced somewhat more harshly than whites, independent of prior criminal history and severity

of current offense. The 32 analyses of State court data reveal that two-thirds of effect sizes indicated that Latinos were sentenced more harshly than whites. The random effects mean odds ratio for the overall measure of unwarranted sentencing disparity was 1.18, which is statistically significant; however, the distribution of Latino/white contrasts was highly variable. Thus, Latinos also were at a statistically small disadvantage in sentencing outcomes in comparison to whites. Once again, however, the influence of being a minority is highly variable.

In regards to sentencing outcomes of Asians and Native Americans, too few contrasts were available for a full meta-analytic synthesis. The findings from the few available studies, however, indicate that sentencing disparity between these racial groups and whites were statistically small. The average difference in sentencing outcomes between Native Americans and whites was very small and not statistically significant. The difference between Asians and whites was not statistically or substantively significant in State courts; however, the difference between these groups was statistically significant but statistically small in analyses of Federal court data. The small number of available effect sizes, however, makes drawing firm conclusions for these analyses tenuous.

Overall, the available empirical evidence supports our first hypothesis, in that the above effect size analyses generally found that minorities were disadvantaged in sentencing outcomes. These effects, however, were typically small in a statistical sense but statistically significant. Moreover, the disadvantage against minorities varied substantially from one context to the next context.

The effect size analyses also generally supported our second hypothesis, as those analyses that measured important variables (i.e., race, offense seriousness, criminal history) precisely and/or controlled for factors known to be related to both race/ethnicity and sentence

severity generated smaller estimates of unwarranted disparity. In fact, the moderator variable analysis of African-American/white contrasts revealed a very strong relationship between effect size and precision of criminal history measure, with those studies using more precise criminal history measures yielding considerably smaller effect sizes than those analyses that controlled for criminal history using only a single dichotomous measure. Similarly, analysis of the Latino/white contrasts indicated that estimates of unwarranted disparity decreased as the number of variables measuring criminal history increased. It was also found that analyses that employed a greater number of control variables, especially analyses including controls for defendant SES, were associated with smaller effect sizes than analyses utilizing fewer controls or omitting measures of defendant SES.

By contrast, the effect size analyses indicated that the findings of studies included in this meta-analysis were not related to characteristics of the research sample. It was hypothesized that differences in sample characteristics (e.g., mean age, proportion female) would account for a share of the variation in effect size. The analyses were unequivocal in their rejection of this hypothesis—none of the coded sample characteristics displayed a meaningful relationship to effect size.

One of the more interesting findings is that there was some evidence that the presence of structured sentencing was associated with smaller estimates of unwarranted sentencing disparity, as we hypothesized (hypothesis four). In particular, analyses examining State court data contrasting African-American sentencing outcomes to those of whites, at the bivariate level found that jurisdictions utilizing structured sentencing had smaller average estimates of unwarranted disparity than jurisdictions without such mechanisms. No such association, however, was found in analyses that compared African-American sentencing outcomes to those

of whites using Federal court data; nor was there evidence that the presence of structured sentencing reduced unwarranted disparity between whites and Latinos. These findings offer a complex set of results, not easily interpretable—making conclusions regarding the veracity of our fourth hypothesis difficult. At best, the evidence in support of our fourth hypothesis is mixed.

CHAPTER 5. DISCUSSION AND CONCLUSION

Purpose and Findings

The issue of racial and ethnic disparity in criminal sentencing has been one of the longest standing research topics in all of criminology. At least 70 years of empirical research has focused on this issue without a clear consensus emerging. The present research found over 300 studies related to this subject; and, over the past 30 years, several noteworthy syntheses of this body of research have been conducted using primarily qualitative, narrative review techniques.

This study departs from these earlier reviews in several important ways. First, this research utilizes meta-analytic techniques that systematically and comprehensively review all available research, published and unpublished, assessing the direct influence of race and ethnicity on sentencing outcomes. We believe that the use of meta-analysis in the present research is an innovative approach to addressing an old research question. One of the most common uses of meta-analysis has been to synthesize a body of research assessing the effectiveness of a certain intervention or a group of interventions (e.g., drug treatment programs). From each study included in these traditional meta-analyses, effect sizes are computed by comparing treated respondents to untreated respondents. Coded features of each study are used to predict the circumstances under which the intervention of interest is most effective—perhaps the intervention is most successful with certain kinds of clients or in certain types of settings. The present use of meta-analysis departs from these traditional meta-analyses by focusing on an unusual type of intervention, criminal sanctioning. Effect sizes were computed by comparing the sentences received by minorities (the treated respondents) to those of whites (the untreated comparison group). Features of each study were coded in an attempt to predict under which

circumstances unwarranted sentencing disparity was more likely to occur. By applying a relatively new approach to address an old problem, we believe that a more accurate view of the cumulative findings of the extant literature has been rendered.

Second, this study reviews research on the influence of both race and ethnicity in sentencing outcomes. Prior syntheses overwhelmingly directed their attention solely towards addressing whether African-Americans were disadvantaged in court outcomes in comparison to whites. The present synthesis extends this focus to include sentencing outcomes of other racial and ethnic minorities.

Third and perhaps most importantly, this review goes beyond simply attempting to address the question of whether unwarranted racial/ethnic disparity exists and also attempts to address the question of *why* studies of sentencing disparity often produce inconsistent findings. We hypothesize that the magnitude of unwarranted racial/ethnic disparity varies systematically with characteristics of the research sample, methodology and contextual setting.

In regards to African-Americans, this study's findings show that African-Americans sentenced in State courts are generally punished more harshly than whites, independent of offense seriousness and prior criminal history. While the influence of race is highly variable, this effect was found to be statistically significant but statistically small. Further, the size of this influence did not vary meaningfully over time—suggesting that unwarranted racial disparity is not confined to only earlier analyses. By contrast, the above effect size analysis indicates that the influence of race in Federal courts was not statistically significant and was statistically small, when all analyses of Federal court sentencing were conducted. More recent analyses of Federal court data, however, reveal that the disadvantage experienced by African-Americans was considerably greater than in earlier analyses of Federal court data, and in these more recent

analyses the influence of race (i.e., being African-American) is sizeable and statistically significant.

The present research also found that the magnitude of the unwarranted sentencing disparity disadvantaging African-Americans varied with several other factors. First, estimates of unwarranted racial disparity varied by type of sentencing outcome. Unwarranted sentencing disparity disadvantaging African-Americans was largest when imprisonment decisions were scrutinized and smaller when length of incarcerative sentence was assessed. While it is certainly possible that unwarranted racial disparity is more prominent in the decision to imprison than in sentence length decisions, such a finding is somewhat perplexing. Why would court officials take a defendant's race into account in imprisonment decisions but not sentence length decisions? Perhaps this apparent contradiction is due to the fact that African-American offenders of marginal seriousness are likely to be sent to prison, whereas similar white offenders are typically sentenced to non-incarcerative sentences (the results of the present meta-analysis provide support for this scenario). These incarcerated marginally serious African-American offenders are likely to receive relatively short terms of incarceration, because their combination of prior criminal history and current offense seriousness lack the severity of other offenders. The implication of this scenario is that when analyses of sentence length are conducted using samples of offenders sentenced to terms of incarceration, these marginal offenders pull down the average sentence length for African-American offenders, in relation to that of white incarcerated offenders. And as a result, analyses of sentence length decisions exhibit relatively smaller estimates of unwarranted racial disparity.

Unwarranted racial disparity was also relatively large in discretionary sentencing decisions. The unstructured nature of these discretionary decisions appears to allow race to

affect sentencing outcomes. That is, whereas imprisonment and sentence length decisions are somewhat constrained by various sentencing structures, discretionary sentencing decisions (by definition) are much less tightly regulated, apparently allowing race to enter the decision-making process and affect sentencing outcomes.

A second important source of systemic variation was type of offense. Specifically, sentencing disparity was largest in cases that examined sentences for drug offenses and smallest in cases involving property crimes. The U.S. has experienced several "moral panics" over drugs in its history. Musto argues that in each of these moral panics over drugs: "use of a particular drug was attributed to an identifiable and threatening minority group. . . [Further,] (t)he belief that drug use threatened to disrupt American social structures militated against moves toward drug toleration, such as legalizing drug use for adults, . . . Public response to these minority-linked drugs differed radically from attitudes towards other drugs with similar potential for harm" (Musto 1973:245). In this most recent moral panic, drugs were identified in the public's imagination with violence and non-white inner city dwellers, particularly African-Americans and Latinos (Tonry 1995). In fact, the relationship between drugs, violence, and disadvantaged inner city minorities has become so intertwined in the mainstream's collective imagination that drug crimes have become a symbol of a non-white urban criminal threat spreading into previously safe, largely suburban, areas (Chiricos 1996). This finding that unwarranted disparity is largest in drug cases supports the hypothesis that drug crimes have become a symbolic racial threat and that the criminal justice system has been mobilized in discriminatory manner to quell this threat.

Third, and perhaps most importantly, several methodological features were important moderators of effect size. In support of Kleck's (1981) arguments, those studies that utilized only a single dichotomous measure of criminal history generated considerably larger estimates of

unwarranted disparity than analyses using more precise measures of this important variable. Likewise, analyses using imprecise measures of offense severity also produced larger estimates of unwarranted racial disparity. Furthermore, some of Wilbank's (1987) arguments also were supported by this research; in that, in line with his arguments, analyses employing more control variables for factors known to be related to both severity of sentence and race, especially SES of defendant, yielded smaller estimates of unwarranted disparity than analyses employing fewer control variables. Yet, even after taking into account these methodological features, race still influenced sentencing outcomes.

It is also important to note which variables failed to exhibit meaningful associations with effect size. Perhaps most salient, is the weak negative association between effect size and presence of structure sentencing in analyses of State data. At the bivariate level, jurisdictions employing structured sentencing displayed smaller average levels of unwarranted racial disparity than jurisdictions without such mechanisms, and this association was marginally statistically significant. This difference, however, was not statistically significant once other factors were taken into consideration. Subsequent multivariate analyses, found that while the presence of structured sentencing continued to have a modest negative relationship to effect size, this relationship was not statistically significant. Similarly, at the bivariate level, Southern jurisdictions displayed larger estimates of unwarranted disparity than other regions of the U.S., but this relationship was not statistically significant after taking into account methodological differences between studies.

While a considerably smaller number of studies analyzed contrasts between Latinos and whites, analyses of these contrasts found that Latinos in both State and Federal courts generally were sentenced more harshly than whites. Similar to the findings of African-American/white

contrasts, the influence of ethnicity (i.e., being Latino in comparison to being white) was highly variable, and statistically small but significant. Also similar to the findings of African-American effect sizes, unwarranted Latino/white sentencing disparity was greatest when imprisonment and discretionary sentencing outcomes were examined and in offenses involving drugs. In contrast to the analyses of African-American/white contrasts, methodological features of the analyses were not as strongly related to effect size. In fact only two methodological characteristics, the use of selection bias corrections and number of variables measuring criminal history, were statistically related to effect size.

The research assessing sentencing practices of Asians and Native Americans was scant (perhaps too few for a full meta-analytic synthesis), but the available research indicates that unwarranted sentencing disparity involving these minority groups is minimal. A possible exception to this general conclusion is the finding that Asians in Federal courts were punished more harshly than whites. More research is needed regarding the sentencing outcomes of these groups before more definitive conclusions can be drawn.

These findings undermine the so-called "no discrimination thesis" which contends that once adequate controls for other factors, especially legal factors (i.e., criminal history and severity of current offense), are controlled unwarranted sentencing disparity disappears. Independent of other factors, minorities were sentenced more harshly than whites on average. The observed differences between whites and minorities generally were small suggesting that discrimination is not the primary cause of the overrepresentation of minorities in U.S. prisons. The extant literature also shows, however, considerable unwarranted racial and ethnic disparities in analyses involving drug offenses, imprisonment decisions, and recently collected Federal data. Furthermore, while the disadvantage suffered by minorities may be small when only one

sentencing episode is considered, the cumulative disadvantage endured by minorities may be considerably larger. Given the strong relationship between prior criminal history and sentencing outcomes, small disadvantages suffered in the past may have substantial effects in subsequent sentencing episodes. Thus, the statistically small minority effects may add up over time to produce considerably larger cumulative effects.

Limitations of Current Research

It is important to keep in mind that this research assessed only the *direct* impact of race/ethnicity on sentencing outcomes. The interactional effects of race could not be synthesized because of the scattered and disparate nature of this research. In other words, there are relatively few studies exploring the interactional relationships between race/ethnicity and other factors on sentencing decisions. Further, what research does exist, does not examine interactions between race/ethnicity and the same set of third factors. These difficulties make a meta-analytic synthesis of these interactional relationships unsuitable at this time.

This is an important limitation of this research, as a growing body of research indicates that minority status when combined with other factors has large influences on sentencing outcomes, independent of other factors (Steffensmeier et al. 1998; Spohn and Holleran 2000; Chiricos and Bales 1991). For example, Chiricos and Bales (1991) found that unwarranted sentencing disparity disadvantaging African-Americans was greatest for unemployed young African-American males; i.e., race interacted with unemployment status and age. Somewhat similarly, Steffensmeier, Ulmer, and Kramer (1998) and Spohn and Holleran (2000) both found that race (and ethnicity in Spohn and Holleran) interacted with age and gender in a manner that produced large disadvantages in imprisonment decisions for young African-American males.

Another limitation of the current research is that it does not consider the potential indirect influences of race/ethnic on sentencing outcomes. Once again, there are too few studies assessing the indirect effects of race/ethnicity, and those studies that do assess such effects do not focus on the same set of third variables. As an example of this line of research, Spohn and DeLone (2000) found that African-Americans had higher odds of being detained during the pretrial phase and defendants who were detained were much more likely to be sentenced to terms of imprisonment than defendants released during pretrial. These authors interpret these findings as being indicative of an indirect race effect. Unfortunately, few other studies report analyses of indirect race/ethnicity effects, and therefore meta-analytic synthesis of such studies does not appear to be a fruitful endeavor at this time.

It needs to re-emphasized that meta-analytic estimates of racial and ethnic disparity reported in this research are directly tied and influenced by the primary author's analyses. If the primary author's models were significantly distorted by model misspecification, and if the coded moderator variables utilized in the present research were not able to discern flawed analyses from rigorous analyses, then the results of this study also will be distorted. Thus, a final potential limitation of this research is its reliance on studies of questionable methodological rigor. However, to the extent that the coded moderator variables were able to discern flawed studies from methodologically rigorous studies, this issue is non-problematic.

Implications of Findings

This research has implications for researchers, theorists, and policy-makers. Perhaps foremost, this research clearly demonstrates the importance of using precise measures of key variables and controlling for third factors related to both race/ethnicity and sentencing outcomes.

Studies that utilized imprecise measures of criminal history and offense seriousness, or omitted key control variables such as defendant SES were associated with systematically larger estimates of unwarranted sentencing disparity than studies using more precise measures. Clearly, these findings suggest that it is imperative that researchers employ numerous, precisely measured variables tapping these important constructs in future research.

The present findings also suggest that researchers need to conduct disaggregated data analyses, as the influence of race and ethnicity varied by type of offense and type of sentencing outcome. Stated differently, the size of the unwarranted racial/ethnic sentencing disparity varied by type of offense and type of sentencing outcome, and therefore future research must continue the trend of conducting disaggregated (or interactional) analyses to detect such variation.

This research also suggests that subsequent research should strongly consider conducting analyses at lower levels of aggregation, as the current meta-analysis found that analyses of highly aggregated data (e.g., analyses of data pooled from all jurisdictions within one state) produced systematically smaller estimates of unwarranted sentencing disparity than analyses conducted at a lower (smaller) levels of aggregation. This finding strongly suggests that aggregation bias affects estimates of unwarranted racial/ethnic disparity in analyses using higher levels of aggregation. Similar findings have been found by earlier researchers (Nelson, 1992; Zimmerman and Frederick, 1984); however, it appears that few researchers have seriously taken into consideration the potentially distorting effects of aggregation bias.

The implications of this research for policy-makers are more reserved. The present research does suggest that policy-makers need to examine the racial and ethnic neutrality of the sentencing policies and practices both generally and especially in regards to certain specific types of sentencing decisions. The persistent finding of differences in sentencing outcomes

between minorities and whites suggests that policy-makers' efforts to achieve racial/ethnic neutrality have not been completely successful in eliminating such disparities. Yet, this research does provide some evidence that structured sentencing mechanisms at the State level are associated with smaller unwarranted sentencing disparities—signifying that these interventions have been at least marginally successful in this regard. Furthermore, the relatively larger sentencing disparities evident in imprisonment decisions and drug offenses suggests that policy-makers need to re-evaluate, and potentially alter, sentencing policies in these arenas.

Conclusion

The results of this meta-analytic synthesis of the race/ethnicity and sentencing research indicates that minorities were sentenced more harshly than whites. The differences in sentencing outcomes between these groups generally were statistically significant but statistically small. Larger estimates of unwarranted sentencing disparity were found in analyses examining imprisonment and discretionary decisions and drug offenses. Smaller estimates of unwarranted sentencing disparity were found in analyses that employed more controls variables, especially those that controlled for defendant SES, and those that utilized precise measures of key variables (prior criminal record and current offense seriousness). However, even when consideration was confined to those analyses employing key controls and precise measures of key variables, statistically significant but statistically small differences in sentencing outcomes persisted. These findings call into question the so-called "no discrimination thesis."

APPENDIX A. CODING FORMS

Race & Sentencing Meta-Analysis
Eligibility Criteria Form

|__|__|__| Study Identification Number

_____ Authors' Last Names

_____ Date of Publication

yes *no*

|__| |__| Does this study:
1) analyze a sentencing outcome in adult courts; non-capital cases;
2) simultaneously control for offense seriousness and prior record?
3) measure the direct influence of race/ethnicity;

Note: If study satisfies the above criterion check "yes"

|__| If NO, what criterion was not met (take first failed criterion)?
1 = Doesn't analyze SENTENCING outcomes in non-capital cases
2 = Doesn't simultaneously control off. seriousness/criminal history
3 = Doesn't measure direct influence of race/ethnicity
4 = Not an empirical study – Relevant literature review;
5 = Other, Specify: _____

yes *no*

|__| |__| (If eligible,) Is enough information reported to code an effect size?

|__| If NO, what information is missing?
1 = No standard deviations, analysis specific N's;
2 = Parameter estimates are not reported
3 = Type of analysis is uncodeable;
4 = Other, Specify: _____

yes *no*

|__| |__| Study analyzes independent data?

If not, which study does the data overlap with? _____

Race & Sentencing Meta-Analysis
Coding Forms

STUDY-LEVEL

1. STUDY ID: _____

2. AUTHORS' LAST NAMES & YEAR: _____

_____

3. PUBLICATION TYPE: _____
   1 = Book;
   2 = Book Chapter;
   3 = Journal (peer reviewed);
   4 = Unpublished or Pseudo-Published (Dissertations, Government Reports)

4. NUMBER OF DIFFERENT SAMPLES/CONTEXTS REPORTED: _____

5. NOTES: _____

_____

_____

SAMPLE/CONTEXT-LEVEL

1. STUDY ID: _____

2. CONTEXT ID: _____

3. JURISDICTION NAME (Name of State/County/etc.): _____

4. JURISDICTION TYPE: _____
   1 = City/County;
   2 = State;
   3 = Federal
   4 = Other

5. REGION OF JURISDICTION: _____
   1 = Southern (i.e., AL, AR, FL, GA, LA, MS, NC, SC, TX, TN, VA)
   2 = Not Southern or Mixture of Southern and other regions.

6. DATA YEAR(s): _____

7a. JURISDICTION HAD STRUCTURED SENTENCING (Y/N): _____
7b. If Yes, check one:
   DETERMINATE ____
   PRESUMPTIVE ____
   VOLUNTARY ____

   Note: If there is no discussion of this issue consult *National Assessment of Structured Sentencing (*BJA, 1996)

8. HOW IS CRIMINAL HISTORY MEASURED?: _____
   1 = One dichotomous measure (e.g., no prior record vs. some prior record)
   2 = Multiple dichotomous measures or ordinal/interval measure

9. HOW IS OFFENSE SEVERITY MEASURED?: _____
   1 = Common law offense type (e.g., violent, property, drug)
   2 = Rating of offense severity (e.g., guideline scale or other scale)
   3 = Both of the above

10. AVERAGE AGE OF SAMPLE: _____

11. PROPORTION OF SAMPLE FEMALE: _____

12. NOTES: _____

_____

OUTCOME LEVEL

1. STUDY ID: _____

2. CONTEXT ID: _____

3. OUTCOME ID: _____

4. OUTCOME LABEL (as used by authors): _____

5. TYPE OF SENTENCING DECISION: _____
   1 = Imprisonment Decision,
   2 =Sentence Length,
   3 = Both,
   4 = Discretionary Lenience,
   5 = Discretionary Harshness

6. TOTAL NUMBER OF VARIABLES IN MODEL: _____

7. NUMBER OF VARIABLES RELATING TO CRIMINAL HISTORY: _____

8. NUMBER OF VARIABLES RELATING TO CURRENT OFFENSE: _____

9. NUMBER OF CONTROL VARIABLES: _____
Note: Omit variable relating to race/ethnicity, criminal history, or offense type.

10. ANALYSIS TAKES SELECTION BIAS INTO ACCOUNT (Y/N): _____

11. CONTROLS FOR TYPE OF COUNSEL (Y/N): ____

12. CONTROLS FOR MODE OF DISPOSITION (Y/N): _____

13. CONTROLS FOR DEFENDANT SES (Y/N): _____
Note: SES can be measured by variances relating to income or employment status

EFFECT SIZE LEVEL

1. STUDY ID: _____

2. CONTEXT ID: _____

3. OUTCOME ID: _____

4. EFFECT SIZE ID: _____

5. RACIAL/ETHNIC CONTRAST: _____
   1 = African American vs. white,
   2 = Non-White vs. white or Non-African-American vs. African-American,
   3 = Hispanic/Latino vs. white,
   4 = Asian vs. white,
   5 = Native American vs. white

6. TYPE OF OFFENSE: _____
   1 = Mixed (mixture of the below types),
   2 = Violent (e.g., assault, rape, robbery),
   3 = Property (e.g., theft, fraud, stolen property)
   4 = Drug (e.g., possession, sales, distribution)
   5 = Other (e.g., weapons, public order offenses)

7. COMPARED TO WHITES, DOES GROUP INDICATING MINORITIES RECEIVE MORE SEVERE OUTCOME (Ignore statistical significance)? (Y/N): __

8. STATISTICALLY SIGNIFICANT DIFFERNECE? (Y/N): _____

9. SAMPLE SIZE FOR THIS EFFECT SIZE: _____

10. TYPE OF EFFECT SIZES (Log Odds Ratio or OLS): _____

11. SD OF DEPENDENT VARIABLE (Only for OLS): _____

12. LOG ODDS RATIO or UNSTANDARDIZED OLS COEFFICIENT: _____

13. NOTES: _____

_____

_____

APPENDIX B. ESTIMATION OF CELL FREQUENCIES OF A 2X2 CONTINGENCY TABLE

The following is an example of the process utilized to estimate the cell frequencies of the implied 2 x 2 contingency table based on a reported odds ratio and marginal frequencies. Allison (1999:12) reports the following 2 x 2 contingency table (the table has been transposed) of death sentence by race of defendant:

	Death	Life	Total
Black	28	45	73
Non-Blacks	22	52	74
Total	50	97	147

In order to demonstrate our estimation process, pretend we had the same contingency table, but we did not know the individual cell frequencies. We only know the odds ratio and marginal frequencies (or equivalently the odds ratio, the total sample size, and the marginal proportions); this information is reported in the vast majority of studies. In this example the odds ratio for this contingency table equals 1.4707 [(28*52)/(45*22)], and the row and column marginals are reproduced below.

	Death	Life	Total
Black	n_{m1}	n_{m0}	73
Non-Blacks	n_{w1}	n_{w0}	74
Total	50	97	147

Inputting these values into the formulas given on page 72, leads to the following values of *a*, *b*, and *c*.

$$a = OR - 1 = 1.4707 - 1 = 0.4707$$
$$b = n_c - n_{r2} - OR(n_c) - OR(n_r) = 50 - 74 - 1.4707(50) - 1.4707(73) = -204.9$$
$$c = OR(n_c)(n_r) = 1.4707(50)(147) = 5368.1$$

Substituting these values into equation 6, yields

$$n_{m1} = \frac{204.9 - \sqrt{(-204.9)^2 - 4(0.4707)(5368.1)}}{2(0.4707)} = 28$$

Once the frequency of this cell has been calculated determining the frequencies of the other cells is just a matter of subtraction. The frequency for n_{m0} equals the row frequency minus n_{m1}; that is, 73-28 = 45. The frequency for nw1 equals the column frequency minus n_{m1}, or 50-28 = 22. Lastly, the value of n_{w0} equals the frequency of the second row minus n_{w1}; that is 74-22 = 52.

It is important to note that when the odds ratio equals 1 the above formula will not work. In this instance, the cell frequencies can be computed by dividing the cross-product of the corresponding row and column marginals by the total sample size.

APPENDIX C. DEPENDENT STUDIES

Albonetti, Celesta A. 1997. "Sentencing Under the Federal Sentencing Guidelines: Effects of Defendant Characteristics, Guilty Pleas, and Departures on Sentence Outcomes for Drug Offenses, 1991-1992." *Law & Society Review* 31(4):789-822.

———. 1998. "Direct and Indirect Effects of Case Complexity, Guilty Pleas, and Offender Characteristics on Sentencing for Offenders Convicted of a White-Collar Offense Prior to Sentencing Guidelines." *Journal of Quantitative Criminology* 14(4):353-78.

Bales, William D. 1987. "Race and Class Effects on Criminal Justice Prosecution and Punishment Decisions: a Test of Some Conflict Theory Propositions." Dissertation, Florida State University, Ann Arbor, MI .

Baxter, Brent L. 1991. "The Contextual Influence of County Racial Characteristics on Racial Disparities in Criminal Sanctions." Dissertation, University of Washington, Ann Arbor, MI.

Brennan, Pauline K. 2002. *Women Sentenced to Jail in New York City*. New York: LFB Scholarly Publishing.

Britt, Chester L. 2000. "Social Context and Racial Disparities in Punishment Decisions." *Justice Quarterly* 17(4):707-32.

Bushway, Shawn and Anne M. Piehl. 2001. "Legal Factors Are Not the Law: An Alternative Approach to the Study of Racial Discrimination in Sentencing in the Era of Sentencing Reform." Unpublished Manuscript.

Bushway, Shawn D. and Anne M. Piehl. 2001. "Judging Judicial Discretion: Legal Factors and Racial Discrimination in Sentencing." *Law & Society Review* 35(4):733-64.

Chen, Huey T. 1981. "Disposition of Felony Arrests: a Sequential Analysis of the Judicial Decision Making Process." Dissertation , University of Massachusetts, Ann Arbor, MI.

Clarke, Stevens H. 1987. *Felony Sentencing in North Carolina, 1976-1986: Effects of Presumptive Sentencing Legislation*. University of North Carolina, Institute of Government.

Clarke, Stevens H. and Gary G. Koch. 1977. *Alaska Felony Sentencing Patterns: A Multivariate Statistical Analysis*. Anchorage, AL: Alaska Judicial Council.

Clayton, Obie Jr. 1983. "A Reconsideration of the Effects of Race in Criminal Sentencing." *Criminal Justice Review* 8(2):15-20.

Coppom, John T. 1978. "Felony Sentencing Patterns of Three Criminal Court Judges: A Regression Model." Dissertation, Ohio State University, Ann Arbor, MI.

Dungworth, Terence. 1978. *An Empirical Assessment of Sentencing Practices in the Superior Court of the District of Columbia*. Washington, D.C.: Institute for Law and Social Research .

Engen, Rodney L. and Sara Steen. 2000. "The Power to Punish: Discretion and Sentencing Reform in the War on Drugs." *American Journal of Sociology* 105(5):1357-95.

Everett, Ronald C. and Roger A. Wojtkiewicz. 2002. "Differences, Disparity, and Race/Ethnicity Bias in Federal Sentencing." *Journal of Quantitative Criminology* 18(2):189-211.

Frazier, Charles E. and Wibur Bock. 1982. "Effects of Court Officials on Sentence Severity." *Criminology* 20(2):257-72.

———. 1964. "Inter- and Intra-Racial Crime Relative to Sentencing." *Journal of Criminal Law, Criminology, and Police Science* 55:348-58.

Griffin, Elizabeth K. 1994. "An Evaluation of Maryland's Sentencing Guidelines: Have They Reduced Disparity in Sentencing?" Thesis, University of Maryland, College Park, MD.

Hagan, John, John D. Hewitt, and Duane Alwin. 1979. "Ceremonial Justice: Crime and Punishment in a Loosely Coupled System." *Social Forces* 58(2):506-27.

Harig, Thomas J. 1990. " The Influence of Extra-Legal Factors in Youthful Offender Adjudications." Dissertation, State University of New York at Albany, Ann Arbor, MI.

Hebert, Christopher G. 1997. "Sentencing Outcomes of Black, Hispanic, and White Males Convicted Under Federal Sentencing Guidelines." *Criminal Justice Review* 22:133-56.

Holmes, Malcolm D., Howard C. Daudistel, and Ronald A. Farrell. 1987. "Determinants of Charge Reductions and Final Dispositions in Cases of Burglary and Robbery." *Journal of Research in Crime and Delinquency* 24(3):233-54.

Holmes, Malcolm D., Harmon M. Hosch, Howard C. Daudistel, Dolores A. Perez, and Joseph B. Graves. 1993. "Judges' Ethnicity and Minority Sentencing: Evidence Concerning Hispanics." *Social Science Quarterly* 74(3):496-506.

Huang, W. S. W., Mary A. Finn, R. B. Ruback, and Robert R. Friedmann. 1996. "Individual and Contextual Influences on Sentence Lengths: Examining Political Conservatism." *Prison Journal* 76(4):398-409.

Humphrey, John and Timothy Fogarty. 1987. "Race and Plea Bargained Outcomes: A Research Note." *Social Forces* 66(1):176-82.

Hutton, C., F. Pommersheim, and S. Feimer. 1996. "'I Fought the Law and the Law Won.'" Pp. 209-20 in *Native Americans, Crime, and Justice*, Editors Marianne O. Nielsen and Robert A. Silverman. Boulder, CO: Westview Press.

Kautt, Paula M. 2002. "Location, Location, Location: Interdistrict and Intercircuit Variation in Sentencing Outcomes for Federal Drug-Trafficking Offenses." *Justice Quarterly* 19(4):633-71.

Kempf-Leonard, Kimberly and Lisa L. Sample. 2001. "Have Federal Sentencing Guidelines Reduced Severity? An Examination of One Circuit." *Journal of Quantitative Criminology* 17(2):111-44.

———. 1990. "Race and Imprisonment Decisions in California." *Science* 247:812-16.

Knapp, Kay A. 1982. "Impact of the Minnesota Sentencing Guidelines on Sentencing Practices." *Hamline Law Review* 5(2):237-56.

Kruttschnitt, Candace. 1982. "Respectable Women and the Law." *Sociological Quarterly* 23(2):221-34.

LaFree, Gary D. 1985. "Official Reactions to Hispanic Defendants in the Southwest." *Journal of Research in Crime and Delinquency* 22(3):213-37.

Litchtenstein, Karen R. 1982. "Extra-Legal Variables Affecting Sentencing Decisions." *Psychological Reports* 50:611-19.

Lotz, Roy and John D. Hewitt. 1977. "The Influence of Legally Irrelevant Factors on Felony Sentencing." *Sociological Inquiry* 47(1):39-48.

Maxfield, Linda D. and John H. Kramer. 1998. *Substantial Assistance: An Empirical Yardstick Gauging Equity in Current Federal Policy and Practice*. Washington, D.C.: U.S. Sentencing Commission.

McCarthy, John P., Niel Sheflin, and Joseph J. Barraco. 1979. *Report on the Sentencing Guidelines Project to the Administrative Director to the Courts: On the Relationship Between Race and Sentencing*. Trenton, NJ: State of New Jersey Administrative Office of the Courts.

Meyer, Jon'a F. 1995. "'Doing Justice' in the Peoples' Courts: Sentencing by Municipal Court Judges." Dissertation, University of California Irvine, Ann Arbor, MI.

Miethe, Terance D. and Charles A. Moore. 1985. "Socioeconomic Disparities Under Determinate Sentence Systems: A Comparison of Preguideline and Postguideline Practices in Minnesota." *Criminology* 23(2):337-63.

———. 1986. "Racial Differences in Criminal Processing: The Consequences of Model Selection on Conclusions About Differential Treatment." *Sociological Quarterly* 27(2):217-37.

Minnesota Sentencing Guidelines Commission. 1984. *The Impact of the Minnesota Sentencing Guidelines: Three Year Evaluation*. St. Paul, MN: Minnesota Sentencing Guidelines Commission.

Moore, Charles A. and Terence D. Miethe. 1986. "Regulated and Unregulated Sentencing Decisions: An Analysis of First-Year Practices Under Minnesota's Felony Sentencing Guidelines." *Law & Society Review* 20(2):253-77.

Mustard, David B. 1997. "Two Essays on the Economics of Crime." Dissertation, University of Chicago, Ann Arbor, MI.

Myers, Martha A. 1987. "Economic Inequality and Discrimination in Sentencing." *Social Forces* 65(3):746-66.

Myers, Martha A. and Susette M. Talarico. 1986. "The Social Contexts of Racial Discrimination in Sentencing." *Social Problems* 33(3):236-51.

Nelson, James F. 1991a. *Racial and Ethnic Disparities in Processing Persons Arrested for Misdemeanor Crimes: New York State, 1985-1986*. Albany, NY: Division of Criminal Justice Services.

———.1991b. "Disparity in the Incarceration of Minorities in New York State, 1985-1986." Pp. 145-60 in *Race and Criminal Justice*, Editors Michael J. Lynch and E. B. Patterson. New York, NY: Harrow and Heston.

———. 1994. "A Dollar or a Day: Sentencing Misdemeanants in New York State." *Journal of Research on Crime and Delinquency* 31(2):183-201.

———. 1995. *Disparities in Processing Felony Arrests in New York State , 1990-1992*. Albany, NY: New York State Division of Criminal Justice Services.

Nobling, Tracy, Cassia Spohn, and Miriam Delone. 1998. "A Tale of Two Counties: Unemployment and Sentence Severity." *Justice Quarterly* 15(3):401-27.

Pope, Carl E. 1975a. *The Judicial Processing of Assault and Burglary Offenders in Selected California Counties*. Washington, D.C.: National Criminal Justice Information and Statistics Service.

Smith, Douglas A. 1986. "The Plea Bargaining Controversy." *Journal-of-Criminal-Law-and-Criminology* 77(3):949-68.

Souryal, Claire and Charles Wellford. 1997. "An Examination of Unwarranted Sentencing Disparity Under Maryland's Voluntary Sentencing Guidelines." Unpublished manuscript. Available at: http://www.gov.state.md.us/sentencing/reports/disparity.html.

Spohn, Cassia. 1990. "The Sentencing Decisions of Black and White Judges: Expected and Unexpected Similarities." *Law and Society Review* 24(5):1197-216.

———. 1992. "An Analysis of the 'Jury Trial Penalty' and Its Effect on Black and White Offenders." *Justice Professional* 7(1):93-112.

———. 1994. "Crime and the Social Control of Blacks: Offender/Victim Race and the Sentencing of Violent Offenders." Pp. 249-68 in *Inequality, Crime, and Social Control*, Editors George S. Bridges and Martha A. Myers. Boulder, CO: Westview Press, Inc.

———. 1999. "Gender and Sentencing of Drug Offenders: Is Chivalry Dead?" *Criminal Justice Policy Review* 9(3 & 4):365-99.

Spohn, Cassia and Dawn Beichner. 2000. "Is Preferential Treatment of Female Offenders a Thing of the Past? A Multisite Study of Gender, Race, and Imprisonment." *Criminal Justice Policy Review* 11(2):149-84.

Spohn, Cassia and Jeffrey Spears. 1996. "The Effect of Offender and Victim Characteristics on Sexual Assault Case Processing Decisions." *Justice Quarterly* 13(4):649-79.

Spohn, Cassia, Susan Welch, and John Gruhl. 1985. "Women Defendants in Court: The Interaction Between Sex and Race in Convicting and Sentencing." *Social Science Quarterly* 66(1):178-85.

Steffensmeier, Darrell and Chester L. Britt. 2001. "Judges' Race and Judicial Decision Making: Do Black Judges Sentence Differently." *Social Science Quarterly* 82(4):749-64.

Steffensmeier, Darrell and Mark Motivans. 2000. "Older Men and Older Women in the Arms of Criminal Law: Offending Patterns and Sentencing Outcomes." *Journal of Gerontology* 55(3):141-51.

Steffensmeier, Darrell, Jeffery Ulmer, and John Kramer. 1998. "The Interaction of Race, Gender, and Age in Criminal Sentencing: The Punishment Cost of Being Young, Black and Male." *Criminology* 36(4):763-97.

Sutton, L. P. 1978. *Variations in Federal Criminal Sentences: A Statistical Assessment at the National Level*. Albany, NY: Criminal Justice Research Center.

Tiffany, Lawrence, Yakov Avichai, and Geoffrey Peters. 1975. "A Statistical Analysis of Sentencing Federal Courts: Defendants Convicted After Trial, 1967-1968." *Journal of Legal Studies* 4(2):369-90.

United States Sentencing Commission. 1995. *Substantial Assistance Departures in the United States Courts*. Washington, D.C.: United States Sentencing Commission.

Unnever, James D. 1982. "Direct and Organizational Discrimination in the Sentencing of Drug Offenders." *Social Problems* 30:212-25.

Unnever, James D. and Larry A. Hembroff. 1988. "The Prediction of Racial/Ethnic Sentencing Disparities: An Expectation States Approach." *Journal of Research in Crime and Delinquency* 25(1):53-82.

Walsh, Anthony. 1987. "The Sexual Stratification Hypothesis and Sexual Assault in Light of the Changing Conceptions of Race." *Criminology* 25(1):153-73.

Welch, Susan, John Gruhl, and Cassia Spohn. 1984. "Sentencing: The Influence of Alternative Measures of Prior Record." *Criminology* 22(2):215-27.

Welch, Susan and Cassia Spohn. 1986. "Evaluating the Impact of Prior Record on Judges' Sentencing Decisions: A Seven-City Comparison." *Justice Quarterly* 3(4):389-407.

Williamson, Harold E. 1980. "Variables Related to Sentencing of Incarcerated Homicide Offenders in Texas." Dissertation, Sam Houston State, Ann Arbor, MI.

Zatz, Marjorie S. 1984 "Race, Ethnicity, and Determinate Sentencing: A New Dimension to an Old Controversy." *Criminology* 22(2):147-71.

APPENDIX D. UNCODEABLE STUDIES

Alvarez, Alexander. 1996. "American Indians and Sentencing Disparity: An Arizona Test." *Journal of Criminal Justice* 24(6):549-61.

Bernstein, Ilene N., William R. Kelly, and Patricia A. Doyle. 1977. "Societal Reaction to Deviants: The Case of Criminal Defendants." *American Sociological Review* 42(5):743-55.

Burke, Peter J. and Austin T. Turk. 1975. "Factors Affecting Postarrest Dispositions: A Model for Analysis ." *Social Problems* 22(3):313-32.

Chiricos, Theodore G. and Gordon P. Waldo. 1975. "Socioeconomic Status and Criminal Sentencing: An Empirical Assessment of a Conflict Perspective." *American Sociological Review* 40(6):753-72.

Clarke, Stevens H. and Gary G. Koch. 1976. "The Influence of Income and Other Factors on Whether Criminal Defendants Go to Prison." *Law & Society Review* 11(1):57-92.

Clarke, Stevens H., Susan T. Kurtz, Elizabeth W. Rubinsky, and Donna J. Schleicher. 1982. *Prosecution and Sentencing in North Carolina*. Chapel Hill, NC: Institute of Government, University of North Carolina.

Engle, Charles D. 1971. "Criminal Justice in the City: A Study of Sentence Severity and Variation in the Philadelphia Criminal Court System." Dissertation , Temple University, Ann Arbor, MI.

Gibson, James L. 1978. " Race As a Determinant of Criminal Sentences: A Methodological Critique and a Case Study ." *Law & Society Review* 12:455-78.

Hutton, Chris, Frank Pommersheim, and Steve Feimer. 1989. "'I Fought the Law and the Law Won': A Report on Women and Disparate Sentencing in South Dakota." *New England Journal of Criminal & Civil Confinement* 15(2):177-202.

Jendrek, Margaret P. 1984. "Sentence Length: Interactions With Race and Court." *Journal of Criminal Justice* 12(6):567-78.

Kelly, Henry E. 1976. "A Comparison of Defense Strategy and Race As Influences in Differential Sentencing." *Criminology* 14(2):241-49.

Kingsnorth, Rodney, John Lopez, Jennifer Wentworth, and Debra Cummings. 1998. "Adult Sexual Assault: The Role of Racial/Ethnic Composition in Prosecution and Sentencing." *Journal of Criminal Justice* 26(5):359-71.

Maroules, Nicholas and Teresa J. White. 1980. *Alaska Felony Sentences: 1976-1979.* Anchorage, AK: Alaska Judicial Council.

Marx, Thomas J. 1980. *The Question of Racial Disparity in Massachusetts Superior Court Sentences*. Boston, MA: Superior Court Department.

San Marco, Louis R. 1979. "Differential Sentencing Patterns of Among Criminal Homicide Offenders in Harris County, Texas." Dissertation, Sam Houston State University, Ann Arbor, MI.

Shane-Dubow, Sandra, Walter F. Smith, and Kim Burns Haralson. 1979. *Felony Sentencing in Wisconsin*. Madison, WI: Wisconsin Center for Public Policy, Inc.

Stecher, Bridget A. and Richard F. Sparks. 1982. "Removing the Effects of Discrimination in Sentencing Guidelines." Pp. 113-29 in *Sentencing Reform: Experiments in Reducing Disparity*, Editor Martin L. Forst. Beverly Hills, CA: Sage.

Thomson, Randall and Matthew Zingraff. 1981. "Detecting Sentence Disparity: Some Problems and Evidence." *American Journal of Sociology* 84(4):869-80.

Townsend, Jules A. 1996. "The Impact of Race on Variation in the Length of Prison Sentences: Empirical Evidence and Policy Implications." Dissertation, University of Missouri-Kansas City, Ann Arbor, MI.

APPENDIX E. INELIGIBLE STUDIES BY INELIGIBILITY REASON

Lack of Simultaneous Controls

Austin, Thomas L. 1985. "Does Where You Live Determine What You Get? A Case Study of Misdemeanant Sentencing." *Journal of Criminology Law & Criminology* 76(2):490-511.

Bachman, Ronet D., Alexander Alvarez, and Craig Perkins. 1996. "The Discriminatory Imposition of the Law: Does It Affect Sentencing Outcomes for American Indians." Pp. 197-208 in *Native Americans, Crime and Justice*, Editors Marianne O. Nielsen and Robert A. Silverman. Boulder, CO: Westview Press.

Barnes, Carole W. and Rodney Kingsnorth. 1996. "Race, Drug, and Criminal Sentencing: Hidden Effects of the Criminal Law." *Journal of Criminal Justice* 24(1):39-55.

Brock, Deon E., Jon Sorensen, and James W. Marquart. 1988. "The Influence of Race on Prison Sentences for Murder in Twentieth-Century Texas." Pp. 211-25 in *Practical Applications for Criminal Justice Statistics*, Editors M. L. Dantzer, Arthur Lurigio, Magnus J. Seng, and James M. Sinacore. Boston, MA: Butterworth-Heinemann.

Bullock, Henry A. 1961. "Significance of the Racial Factor in the Length of Prison Sentence." *Journal of Criminal Law, Criminology and Police Science* 52:411-17.

Bynum, Timothy S. and Raymond Paternoster. 1984. "Discrimination Revisited: An Exploration of Front Stage and Back Stage Criminal Justice Decision Making." *Sociology and Social Research* 69(1):90-108.

California Administrative Office of the Courts Research and Planning Unit. 2002. *Report to the Legislature Pursuant to Penal Code Section 1170.45: The Disposition of Criminal Cases According to the Race and Ethnicity of the Defendant*. Sacramento, CA: Administrative Office of the Courts Research and Planning Unit.

Chiricos, Theodore G., Phillip D. Jackson, and Gordon P. Waldo. 1972. "Inequality in the Imposition of a Criminal Label." *Social Problems* 19(4):553-72.

Corbo, Cynthia A., Edward Coyle, and Ellen Osbourne. 1990. *Mandatory Sentences for Firearms Offenses in New Jersey*. Newark, NJ: Data Committee, Criminal Disposition Commission.

Crail-Rugotzke, Donna. 1999. "A Matter of Guilt: the Treatment of Hispanic Inmates by New Mexico Courts and the New Mexico Territorial Prison, 1890-1912." *New Mexico Historical Review* 74(3):295-317.

Douglass, Lynne A. 1979. "Justice Midwestern Style: A Study of the Decision Making Process in a County Criminal Justice System." Dissertation, University of Notre Dame , Ann Arbor, MI.

Epstein, Batsheva S. 1981. "Patterns of Sentencing and Their Implementation in Philadelphia City and County, 1795-1829." Dissertation, University of Pennsylvania, Ann Arbor, MI.

Hanbury, Barbara M. 2000. "Are Judges Downwardly Departing From the United States Sentencing Guidelines Based on the Offenders' Personal Characteristics?" Dissertation, University of Maryland, Ann Arbor, MI.

Heaney, Gerald W. 1992. "Revisiting Disparity: Debating Guidelines Sentencing." *American Criminal Law Review* 29(3):771-93.

Kulig, Frank. 1975. "Plea Bargaining, Probation, and Other Aspects of Conviction and Sentencing." *Creighton Law Review* 8:938-54.

Levin, Martin A. 1972. " Urban Politics and Policy Outcomes: The Criminal Courts." Pp. 330-363 in *Criminal Justice: Law and Politics*, Editor George F. Cole. North Scituate, MA: Duxbury Press.

Mann, Coramae R. 1984. " Race and Sentencing of Female Felons: A Field Study." *International Journal of Women's Studies* 7(2):160-172.

Meierhoefer, Barbara S. 1992. *The General Effect of Mandatory Minimum Prison Terms: a Longitudinal Study of Federal Sentences Imposed.* Washington, D.C.: Federal Judicial Center.

Motivans, Mark A. 2001. "Race/Ethnicity Imbalance in Pennsylvania's Prisons: Comparisons Across Case-Processing Stages and White-Black-Hispanic Subgroups." Dissertation, Pennslyvania State University, Ann Arbor, MI.

Munoz, Ed A., David A. Lopez, and Eric Stewart. 1998. "Misdemeanor Sentencing Decisions: The Cumulative Disadvantage Effect of "Gringo Justice"." *Hispanic Journal of Behavioral Sciences* 20(3):298-319.

Munoz, Eddie A. 1999. *Racial Disparities in Imprisonment Rates in Nebraska: A Case Study of Panhandle County*. East Lansing, MI : The Julian Samora Research Institute, Michigan State University.

Nienstedt, Barbara C., Marjorie S. Zatz, and Thomas Epperlein. 1988. "Court Processing and Sentencing of Drinking Drivers: Using New Methodologies." *Journal of Quantitative Criminology* 4(1):39-59.

Perry, Ronald W. 1977. " The Justice System and Sentencing: The Importance of Race in the Military." *Criminology* 15(2):225-34.

Pommersheim, Frank and Steve Wise. 1989. "Going to the Penitentiary: A Study of Disparate Sentencing in South Dakota." *Criminal Justice and Behavior* 16(2):155-65.

Pope, Carl E. 1976. "Influence of Social and Legal Factors on Sentence Dispositions: A Preliminary Analysis of Offender-Based Transaction Statistics." *Journal of Criminal Justice* 4(3):206-21.

———. 1978. "Sentence Dispositions Accorded Assault and Burglary Offenders: an Exploratory Study in Twelve California Counties." *Journal of Criminal Justice* 6(2):151-65.

Rhodes, William M. 1977. "A Study of Sentencing in the Hennepin County and Ramsey County District Courts." *Journal of Legal Studies* 6:333-53.

Sellin, Thorsten. 1935. "Race Prejudice in the Administration of Justice ." *American Journal of Sociology* 41(2):212-17.

Sidanius, Jim. 1988. "Race and Sentence Severity: The Case of American Justice." *Journal of Black Studies* 18(3):273-81.

Smith, Brent L. and Kelly R. Damphouse. 1996. "Punishing Political Offenders: The Effect of Political Motive on Federal Sentencing Decisions." *Criminology* 34(3):289-321.

Smith, Brent L. and Edward H. Stevens. 1984. "Sentence Disparity and the Judge-Jury Sentencing Debate: an Analysis of Robbery Sentences in Six Southern States." *Criminal Justice Review* 9(1):1-7.

Southern Regional Council. 1969. *Race Makes the Difference.* Atlanta, GA: Southern Regional Council.

Uhlman, Thomas. 1978. "Black Elite Decision Making: The Case of Trial Judges." *American Journal of Political Science* 22(4):884-95.

Williams, Oliver and Richard J. Richardson. 1976. "The Impact of Criminal Justice Policy on Blacks in Trial Courts of a Southern State." Pp. 68-84 in *Criminal Justice: Law and Politics,* Second ed. Editor George F. Cole. North Scituate, MA: Duxbury.

Zalman, Marvin, Charles W. Ostrom Jr., Phillip Guilliams, and Garrett Peaslee. 1979. *Sentencing in Michigan: Report of the Michigan Felony Sentencing Project*. Lansing, MI.

Zatz, Marjorie S. 1985. "Pleas, Priors, and Prison: Racial/Ethnic Differences in Sentencing." *Social Science Research* 14:169-93.

Zatz, Marjorie S. and John Hagan. 1985. "Crime, Time, and Punishment: An Exploration of Selection Bias in Sentencing Research." *Journal of Quantitative Criminology* 1(1):103-26.

Didn't Analyze One of the Five Specified Sentencing Outcomes

Albonetti, Celesta A. 1990. "Race and the Probability of Pleading Guilty." *Journal of Quantitative Criminology* 6(3):315-34.

Albonetti, Celesta A. and John R. Hepburn. 1996. "Prosecutorial Discretion to Defer Criminalization: The Effects of Defendant's Ascribed and Achieved Status Characteristics." *Journal of Quantitative Criminology* 12(1):63-81.

Bridges, George S. and Robert D. Crutchfield. 1986. *Racial and Ethnic Disparities in Imprisonment: Final Report*. Seattle, WA: Institute for Public Policy and Management, Graduate School of Public Affairs, University of Washington.

Bridges, George S., Robert D. Crutchfield, and Edith S. Simpson. 1987. "Crime, Social Structure and Criminal Punishment: White and Nonwhite Rates of Imprisonment." *Social Problems* 34(4):345-61.

Brownsberger, William N. 2000. "Race Matters: Disproportionality of Incarceration for Drug Dealing in Massachusetts." *Journal of Drug Issues* 30(2):345-74.

D'Esposito, Julian C. Jr. 1969. "Sentencing Disparity: Causes and Cures." *Journal of Criminal Law, Criminology and Police Science* 60(2):182-94.

Farrell, Ronald A. and Victoria L. Swigert. 1978. "Prior Offense Record As a Self-Fulfilling Prophecy." *Law & Society Review* 12(3):437-53.

Grant, David M. 1983. "An Empirical Examination of Michigan's Felony Firearm Statute." Thesis, Michigan State University, 1983.

Hutner, Michael. 1978. " Sentencing in Massachusetts Superior Court." Dissertation, Syracuse University, Ann Arbor, MI.

Keil, Thomas J. and Gennaro F. Vito. 1995. "Race and the Death Penalty in Kentucky Murder Trials: 1976-1991." *American Journal of Criminal Justice* 20(1):17-36.

LaCasse, Chantel and A. A. Payne. 1999. "Federal Sentencing Guidelines and Mandatory Minimum Sentences: Do Defendants Bargain in the Shadow of the Judge." *Law and Economics* 42:245-69.

Meeker, James W., Paul Jesilow, and Joseph Aranda. 1992. "Bias in Sentencing: A Preliminary Analysis of Community Service Sentences." *Behavior Sciences and the Law* 10(2):197-206.

Myers, Martha A. 1990. " Economic Threat and Racial Disparities in Incarceration: The Case of Postbellum Georgia." *Criminology* 28(4):627-56.

Myers, Samuel L. 1985. "Statistical Tests of Discrimination in Punishment." *Journal of Quantitative Criminology* 1(2):191-218.

Rhodes, William M. 1976. "The Economics of Criminal Courts: A Theoretical and Empirical Investigation ." *Journal of Legal Studies* 5:311-40.

Vining, Aidan R. 1980. " Limits of Individualization: Sentencing Variation and Disparity in California." Dissertation, University of California, Berkeley, Ann Arbor, MI.

———. 1983. "Developing Aggregate Measures of Disparity." *Criminology* 21(2):233-52.

Walsh, Anthony. 1985. "Ideology and Arithmetic: The Hidden Agenda of Sentencing Guidelines." *Journal of Crime and Justice* 8:41-63.

Wolfgang, Marvin E. and Marc Reidel. 1973. "Race, Judicial Discretion, and the Death Penalty." *The Annals of the American Academy of Political and Social Science* 407:119-33.

Yates, Jeff. 1997. "Racial Incarceration Disparity Among States." *Social Science Quarterly* 78(4):1001-10.

No Empirical Analyses

Bateman, Casey, Greg Jones, Kristin Quayle, and Mary Uhlfelder. 2001. *Issues in Maryland Sentencing: An Analysis of Unwarranted Sentencing Disparity*. College Park, MD: Maryland State Commission on Criminal Sentencing Policy.

Black, Donald. 1976. *The Behavior of Law*. San Diego, CA: Academic Press.

Blalock, Hubert. 1967. *Toward a Theory of Minority-Group Relations*. New York, NY: Wiley .

Blumstein, Alfred, Jacqueline Cohen, Susan E. Martin, and Michael H. Tonry, Editors. 1983. *Research on Sentencing: The Search for Reform*, vol. 1Washington, D.C.: National Academy Press.

Bureau of Justice Statistics. 1997. *Lifetime Likelihood of Going to State or Federal Prison*. Washington, DC: U.S. Department of Justice.

———. 2000. *State Court Sentencing of Convicted Felons, 1996*. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Chiricos, Ted and Sarah Eschholz. 2002. "The Racial and Ethnic Typification of Crime and the Criminal Typification of Race and Ethnicity in Local Television News." *Journal of Research in Crime and Delinquency* 39(4):400-420.

Chiricos, Theodore G. and Charles Crawford. 1995. "Race and Imprisonment: A Contextual Assessment of the Evidence." Pp. 281-309 in *Ethnicity, Race, and Crime: Perspectives Across Time and Place*, Editor Darnell F. Hawkins. Albany, NY: State University of New York Press.

Clarke, Stevens H. 1984. "North Carolina's Determinate Sentencing Legislation." *Judicature* 68(4-5):141-52.

Daly, Kathleen and Michael Tonry. 1997. "Gender, Race, and Sentencing." Pp. 201-52 in *Crime and Justice: A Review of Research*, vol. 22, Editor Michael Tonry. Chicago, IL: University of Chicago.

DeLisi, Matt and Bob Regoli. 1999. "Race, Conventional Crime, and Criminal Justice: The Declining Importance of Skin Color." *Journal of Criminal Justice* 27(6):549-57.

Durose, Matthew R., David J. Levin, and Patrick A. Langan. 2001. *Felony Sentences in State Courts, 1998*. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Gottfredson, Don M. 1999. "Measuring Justice: Unpopular Views on Sentencing Theory." Pp. 247-81 in *The Criminology of Criminal Law: Advances in Criminological Theory,* vol. 8, Editors William S. Laufer and Freda Adler. New Brunswick, NJ: Transaction Publishers.

Hagan, John. 1974. "Extra-Legal Attributes and Criminal Sentencing: An Assessment of a Sociological Viewpoint." *Law & Society Review* 8:357-83.

Hagan, John and Kristin Bumiller. 1983. "Making Sense of Sentencing: A Review and Critique of Sentencing Research." Pp. 1-54 in *Research on Sentencing: The Search for Reform,* vol. 2, Editors Alfred Blumstein, Jacqueline Cohen, Susan E. Martin, and Michael H. Tonry. Washington, D.C.: National Academy Press.

Hawkins, Darnell F. 1987. "Beyond Anomalies: Rethinking the Conflict Perspective on Race and Criminal Punishment." *Social Forces* 65(3):719-45.

Kleck, Gary. 1981. "Racial Discrimination in Criminal Sentencing: A Critical Evaluation of the Evidence With Additional Evidence on the Death Penalty." *American Sociological Review* 46(6):783-805.

———. 1985. "Life Support for Ailing Hypotheses: Modes of Summarizing the Evidence for Racial Discrimination in Sentencing." *Law and Human Behavior* 9(3):271-85.

Liska, Allen E. 1987. "A Critical Examination of Macro Perspectives on Crime Control." *Annual Review of Sociology* 13:67-88.

Liska, Allen E. and Mitchell B. Chamlin. 1984. "Social Structure and Crime Control among Macrosocial Units." *American Journal of Sociology* 90(2):383-95.

Lubitz, Robin L. and Thomas W. Ross. 2001. *Sentencing Guidelines: Reflections on the Future*. Washington, D.C.: U.S. Department of Justice, National Institute of Justice.

Mann, Charles C. 1994. "Can Meta-Analysis Make Policy?" *Science* 266(5187):960-962.

Mann, Coramae R. 1993. *Unequal Justice: A Question of Color*. Bloomington, IN: University of Indiana.

Mileski, Maureen. 1971. "Courtroom Encounters: An Observation of a Lower Criminal Court." *Law & Society Review* 5(4):473-538.

Petersilia, Joan and Susan Turner. 1987. "Guideline-Based Justice: Prediction and Racial Minorities." Pp. 151-81 in *Prediction and Classification: Criminal Justice Decision Making*, Editors Don M. Gottfredson and Michael Tonry. Chicago, IL: University of Chicago Press.

Pratt, Travis C. 1998. " Race and Sentencing: A Meta-Analysis of Conflicting Empirical Research Results." *Journal of Criminal Justice* 26(6):513-23.

Sampson, Robert J. and John H. Laub. 1993. "Structural Variations in Juvenile Court Processing: Inequality, the Underclass, and Social Control." *Law & Society Review* 27(2):285-311.

Sampson, Robert J. and Janet L. Lauritsen. 1997. "Racial and Ethnic Disparities in Crime and Criminal Justice in the United States." Pp. 311-74 in *Ethnicity, Crime, and Immigration: Comparative and Cross-National Perspectives*, vol. 27, Editor Michael Tonry. Chicago, IL: University of Chicago Press.

Savelsberg, Joachim J. 1992. "Law That Does Not Fit Society: Sentencing Guidelines As a Neoclassical Reaction to the Dilemmas of Substanitivized Law." *American Journal of Sociology* 97(5):1346-81.

Shane-Dubow, Sandra. 1998. "Introduction to Models of Sentencing Reform in the United States." *Law & Policy* 20(3):231-45.

Spohn, Cassia. 1995. "Courts, Sentences, and Prisons." *Daedalus* 124(1):119-41.

———. 2000. "Thirty Years of Sentencing Reform: The Quest for a Racially Neutral Sentencing Process." Pp. 427-501 in *Policies, Processes, and Decisions of the Criminal Justice System: Criminal Justice 2000*, vol. 3, Editor Julie Horney. Washington, D.C.: National Institute of Justice .

Steffensmeier, Darrell J. 1980. "Assessing the Impact of the Women's Movement on Sex-Based Differences in the Handling of Adult Criminal Defendants." *Crime and Delinquency* 26(3):344-57.

Stern, Barry J. 1985. "Presumptive Sentencing in Alaska." *Alaska Law Review* 2(2):227-69.

Tittle, Charles R. and Debra A. Curran. 1988. "Contingencies for Dispositional Disparities in Juvenile Justice." *Social Forces* 67(1):23-58.

Vincent, Barbara S. and Paul J. Hofer. 1994. *The Consequences of Mandatory Minimum Prison Terms: A Summary of Recent Findings*. Washington, D.C..

Warner, Darren E. 200. " Race and Ethnic Bias in Sentencing Decisions: A Review and Critique of the Literature." Pp. 171-80 in *The System in Black and White: Exploring the Connections Between Race, Crime, and Justice*, Editors Michael W. Markowitz and Delores D. Jones-Brown. Westport, CT: Praeger.

Webb, G. L. 1979. *Conviction and Sentencing: Deception and Racial Discrimination*. Austin, TX: Coker Press.

Wilbanks, William. 1987. *The Myth of a Racist Criminal Justice System*. Monterey, CA: Brooks/Cole.

Zatz, Marjorie S. 1984 "Race, Ethnicity, and Determinate Sentencing: A New Dimension to an Old Controversy." *Criminology* 22(2):147-71.

No Direct Race Measure

Bridges, George S. and Robert D. Crutchfield. 1986. *Racial and Ethnic Disparities in Imprisonment: Final Report*. Seattle, WA: Institute for Public Policy and Management, Graduate School of Public Affairs, University of Washington.

Clancy, Kevin, John Bartolomeo, David Richardson , and Charles Wellford. 1981. "Sentence Decisionmaking: The Logic of Sentence Decisions and the Extent and Sources of Sentence Disparity." *Journal of Criminal Law and Criminology* 72(2):524-54.

Clarke, Stevens H. 1983. "North Carolina's Fair Sentencing Act: What Have the Results Been?" *Popular Government* 49(2):11-16, 40.

Farnworth, Margaret and Patrick M. Horan. 1980. "Separate Justice: An Analysis of Race Differences in Court Processes." *Social Science Research* 9:381-99.

Farrar-Owens, Meredith L. 2001. "The Long Term Impact of Sentencing Guidelines on Unwarranted Disparity in Virginia." Thesis , Virginia Commonwealth University, Ann Arbor, MI.

Flavin, Jeanne. 2001. "Of Punishment and Parenthood: Family-Based Social Control and the Sentencing of Black Drug Offenders." *Gender and Society* 15(4):611-33.

Gilkison, Margaret E. 1985. "Sentencing and Disparity: A Model of Judicial Decision-Making." Dissertation, Michigan State University, Ann Arbor, MI.

Griswold, David. 1987. " Deviation From Sentencing Guidelines: The Issue of Unwarranted Disparity." *Journal of Criminal Justice* 15(4):317-29.

Hofer, Paul J., Kevin R. Blackwell, and R. B. Ruback. 1999. "The Effect of the Federal Sentencing Guidelines on Inter-Judge Sentencing Disparity." *Journal of Criminal Law & Criminology* 90(1):239-322.

Kautt, Paula M. and Cassia Spohn. 2002. "Crack-Ing Down on Black Drug Offenders? Testing for Interactions Among Offenders' Race, Drug Type, and Sentencing Strategy in Federal Drug Sentences." *Justice Quarterly* 19(1):1-35.

LaFree, Gary D. 1980. "The Effect of Sexual Stratification by Race on Official Reactions to Rape." *American Sociological Review* 45(5):842-54.

Landes, William. 1974. " Legality and Reality: Some Evidence on Criminal Procedure." *Journal of Legal Studies* 3(2):287-337.

Myers, Laura B. and Sue T. Reid. 1995. "The Importance of County Context in the Measurement of Sentence Disparity." *Journal of Criminal Justice* 23(3):223-41.

Myers, Martha A. 1979. "Offended Parties and Official Reactions: Victims and the Sentencing of Criminal Defendants." *Sociological Quarterly* 20:529-40.

Nardulli, Peter. 1978. *The Courtroom Elite: An Organizational Perspective on Criminal Justice*. New York, NY: Ballinger.

Payne, A. A. 1997. "Does Inter-Judge Disparity Really Matter? An Analysis of the Effects of Sentencing Reforms in Three Federal District Courts." *International Review of Law and Economics* 17:346-57.

Rhodes, William. 1991. " Federal Criminal Sentencing: Some Measurement Issues With Application to Pre-Guideline Sentencing Disparity." *Journal of Criminal Law & Criminology* 81(4):1002-33.

Roy, Majorie B., Lauren Herbert, and R. L. Copeland. 1980. *Arson in Massachusetts: Sentencing Patterns 1978-1978*. Boston, MA: Massachusetts Commissioner of Probation.

Roy, Marjorie B. and Michelle M. Strout. 1980. *Car Thieves: Sentencing Patterns in Massachusetts 1975-1978*. Boston, MA: Massachusetts Commissioner of Probation.

Sparks, Richard F. and Bridget A. Stecher. 1979. *The New Jersey Sentencing Guidelines: An Unauthorized Analysis*. Newark, NJ: Rutgers School of Criminal Justice.

Truitt, Linda, William M. Rhodes, R. N. Kling, and Q. E. McMullen. 2000. *Multi-Site Evaluation of Sentencing Guidelines: Florida and North Carolina*. Cambridge, MA: Abt Associates Inc.

Vera Institute of Justice. 1981. *Felony Arrests: Their Prosecution and Disposition in New York City's Courts*. New York, NY : Longman Inc.

Welch, Susan, Michael Combs, and John Gruhl. 1988. "Do Black Judges Make a Difference?" *American Journal of Political Science* 32(1):126-36.

Wilkins, Leslie T., Jack M. Kress, Don M. Gottfredson, Joseph C. Calpin, and Arthur M. Gelman. 1978. *Sentencing Guidelines: Structuring Judicial Discretion*. Criminal Justice Research Center.

Willick, Daniel H., Gretchen Gehlker, and Anita M. Watts. 1975. "Social Class As a Factor Affecting Judicial Disposition of Defendants Charged With Criminal Homosexual Acts." *Criminology* 13(1):57-77.

Other Reason

Brantingham, Patricia L. 1985. "Sentencing Disparity: An Analysis of Judicial Consistency." *Journal of Quantitative Criminology* 1(3):281-305.

Sellin, Thorsten. 1928. "The Negro Criminal: A Statistical Note." *The Annals of the American Academy of Political and Social Science* 140:52-64.

APPENDIX F. COMPLETE FOREST PLOT OF AFRICAN-AMERICAN/WHITE CONTRASTS FROM STATE DATA

Figure 1a. African-American/White Contrasts: Forest Plot—State Level Only

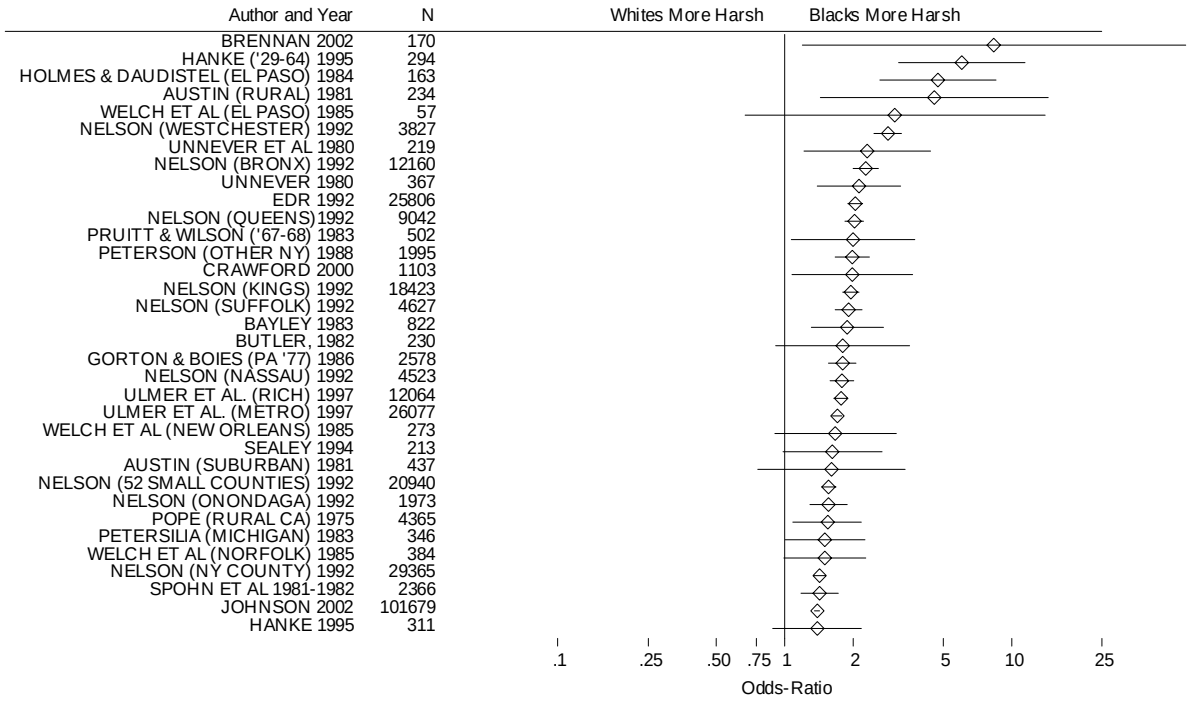


Figure 1a. (Continued) African-American/White Contrasts: Forest Plot—State Level Only

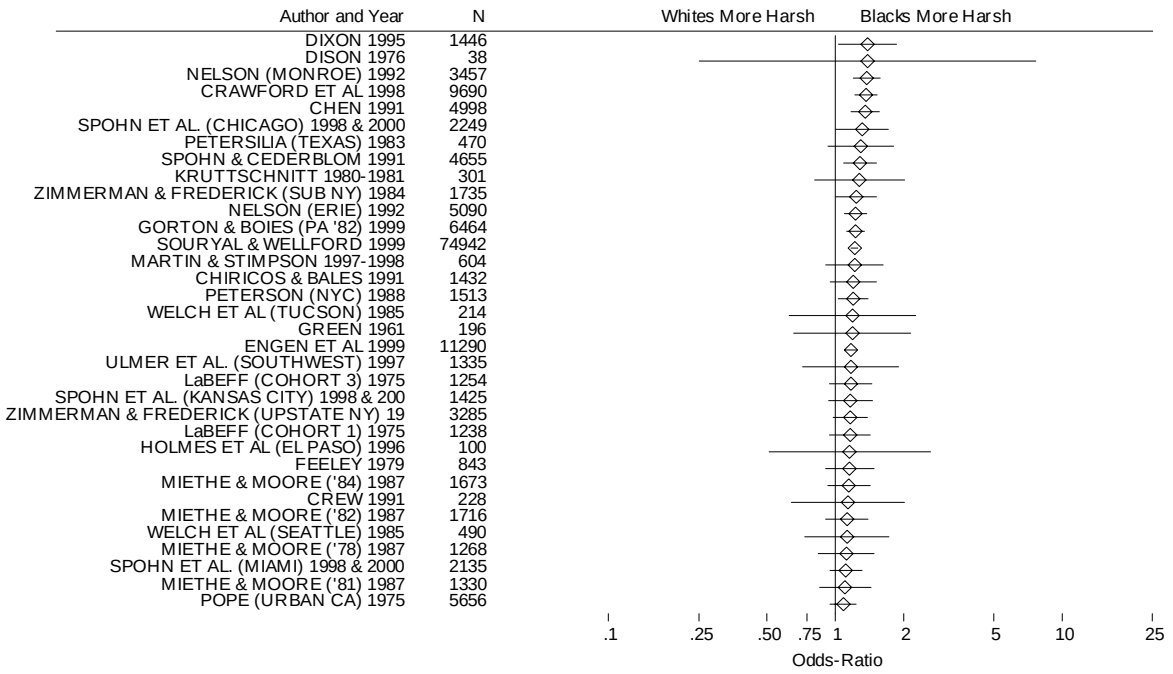
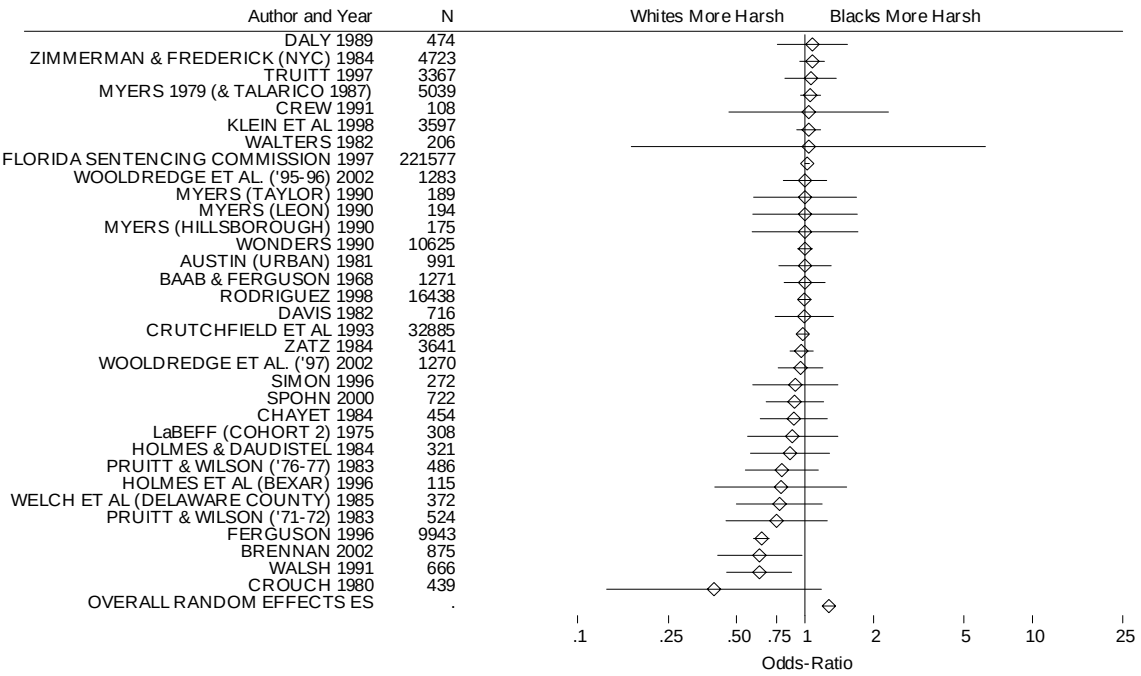


Figure 1a. (Continued) African-American/White Contrasts: Forest Plot—State Level Only



REFERENCE LIST

Albonetti, Celesta A. 1990. "Race and the Probability of Pleading Guilty." *Journal of Quantitative Criminology* 6(3):315-34.

*———. 1991. "An Integration of Theories to Explain Judicial Discretion." *Social Problems* 38(2):247-66.

*———. 1994. "The Symbolic Punishment of White-Collar Offenders." Pp. 269-82 in *Inequality, Crime, and Social Control*, Editors George S. Bridges and Martha A. Myers. Boulder, CO: Westview Press.

———. 1997. "Sentencing Under the Federal Sentencing Guidelines: Effects of Defendant Characteristics, Guilty Pleas, and Departures on Sentence Outcomes for Drug Offenses, 1991-1992." *Law & Society Review* 31(4):789-822.

Aldrich, John H. and Forrest D. Nelson. 1984. *Linear Probability, Logit, and Probit Models*. Beverly Hills, CA: Sage.

Allison, Paul D. 1999. *Logistic Regression Using the SAS System: Theory and Application*. Cary, N.C.: SAS Institute.

Alvarez, Alexander. 1996. "American Indians and Sentencing Disparity: An Arizona Test." *Journal of Criminal Justice* 24(6):549-61.

*Austin, Thomas L. 1981. "The Influence of Court Location on Type of Criminal Sentence: The Rural-Urban Factor." *Journal of Criminal Justice* 9(6):305-16.

———. 1985. "Does Where You Live Determine What You Get? A Case Study of Misdemeanant Sentencing." *Journal of Criminology Law & Criminology* 76(2):490-511.

*Baab, G. A. and W. R. Furgeson Jr. 1968. "Texas Sentencing Practices: A Statistical Study." *Texas Law Review* 45:471-503.

Baxter, Brent L. 1991. "The Contextual Influence of County Racial Characteristics on Racial Disparities in Criminal Sanctions." Dissertation, University of Washington, Ann Arbor, MI.

*Bayley, Kenneth B. 1983. "A Statistical Analysis of Discrimination in the Alaska Criminal Justice System." Dissertation, Claremont University, Ann Arbor, MI.

Becker, Howard S. 1963. *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.

Belenko, Steven, Jeffrey Fagan, and Ko-Lin Chin. 1991. "Criminal Justice Responses to Crack." *Journal of Research in Crime and Delinquency* 28(1):55-74.

Bernstein, Ilene N., William R. Kelly, and Patricia A. Doyle. 1977. "Societal Reaction to Deviants: The Case of Criminal Defendants." *American Sociological Review* 42(5):743-55.

*Bickle, Gayle S. and Ruth D. Peterson. 1991. "The Impact of Gender-Based Family Roles in Criminal Sentencing." *Social Problems* 38(3):372-94.

Blumstein, Alfred, Jacqueline Cohen, Susan E. Martin, and Michael H. Tonry, Editors. 1983. *Research on Sentencing: The Search for Reform*, vol. 1Washington, D.C.: National Academy Press.

*Brennan, Pauline K. 1999. "The Likelihood of Incarceration in New York City: An Analysis of Cases for Black, White, and Hispanic Females Sentenced at Criminal Court Arraignment." Dissertation, State University of New York Albany, Ann Arbor, MI.

*———. 2002. “Sentences Given to Convicted Drug Offenders: An Examination of the Race Effect, by Specific Drug Offenses." *American Society of Criminology Meeting* (Chicago, IL.)

Britt, Chester L. 2000. "Social Context and Racial Disparities in Punishment Decisions." *Justice Quarterly* 17(4):707-32.

Bushway, Shawn D. and Anne M. Piehl. 2001. "Judging Judicial Discretion: Legal Factors and Racial Discrimination in Sentencing." *Law & Society Review* 35(4):733-64.

Callaham, Michael L., Robert L. Wears, Ellen J. Weber, Christopher Barton, and Gary Young. 1998. "Positive-Outcome Bias and Other Limitations in the Outcome of Research Abstracts Submitted to a Scientific Meeting." *Journal of the American Medical Association* 280(3):254-57.

*Chanhatasilpa, Chanchalat. 2000. "Race and Sentencing Revisited: The Sanctioning of Asian American Defendants Under the Federal Sentencing Guidelines." Dissertation, University of Maryland, Ann Arbor, MI.

*Chayet, Ellen F. 1984. "The Influence of Discretion in Post-Conviction Sentencing Disparity." Dissertation, Brandeis University, Ann Arbor, MI.

*Chen, Huey T. 1991. "Dropping in and Dropping Out: Judicial Decisionmaking in the Disposition of Felony Arrests." *Journal of Criminal Justice* 19(1):1-17.

Chiricos, Theodore. 1996. "Moral Panic as Ideology: Race, Drugs, Violence and Punishment in America." Pp. 19-49 in *Justice with Prejudice: Race and Justice in America*, Editors Michael J. Lynch and E. B. Patterson. Albany, NY: Harrow and Heston.

*Chiricos, Theodore G. and William D. Bales. 1991. "Unemployment and Punishment: An Empirical Assessment." *Criminology* 29(4):701-24.

Chiricos, Theodore G. and Charles Crawford. 1995. "Race and Imprisonment: A Contextual Assessment of the Evidence." Pp. 281-309 in *Ethnicity, Race, and Crime: Perspectives across Time and Place*, Editor Darnell F. Hawkins. Albany, NY: State University of New York Press.

Chiricos, Theodore G. and Gordon P. Waldo. 1975. "Socioeconomic Status and Criminal Sentencing: An Empirical Assessment of a Conflict Perspective." *American Sociological Review* 40(6):753-72.

Cooper, Harris and Larry V. Hedges. 1984. "Research Synthesis as Scientific Enterprise." Pp. 3-14 in *The Handbook of Research Synthesis*, Editors Harris Cooper and Larry V. Hedges. New York, NY: Russell Sage Foundation.

Cooper, Harris M. and Robert Rosenthal. 1980. "Statistical Versus Traditional Procedures for Summarizing Research Findings." *Psychological Bulletin* 87(3):442-49.

*Crawford, Charles. 2000. "Gender, Race, and Habitual Offender Sentencing in Florida." *Criminology* 38(1):263-80.

*Crawford, Charles, Ted Chiricos, and Gary Kleck. 1998. "Race, Racial Threat, and Sentencing of Habitual Offenders." *Criminology* 36(3):481-511.

*Crew, B. K. 1991. "Race Differences in Felony Charging and Sentencing: Toward an Integration of Decision-Making and Negotiation Models." *Journal of Crime and Justice* 14(1):99-122.

*Crouch, William H. 1980. "Legal and Extra-Legal Factors Affecting Court Disposition of Felony Offenders." Dissertation, University of Northern Colorado, Ann Arbor, MI.

*Crutchfield, Robert D., Joseph G. Weis, Rodney L. Engen, and Randy R. Gainey. 1993. *Racial/Ethnic Disparities and Exceptional Sentences in Washington State*. Olympia, WA: Washington State Minority and Justice Commission.

*Daly, Kathleen. 1989. "Neither Conflict nor Labeling nor Paternalism Will Suffice: Intersections of Race, Ethnicity, Gender and Family in Criminal Court Decisions." *Crime and Delinquency* 35(1):136-68.

Daly, Kathleen and Michael Tonry. 1997. "Gender, Race, and Sentencing." Pp. 201-52 in *Crime and Justice: A Review of Research*, vol. 22, Editor Michael Tonry. Chicago, IL: University of Chicago.

*Davis, James R. 1982. *The Sentencing Dispositions of New York City Lower Court Criminal Judges*. Washington, D.C.: University Press of America.

*Dison, Jack E. 1977. "Empirical Examination of Conflict Theory: Race and Sentence Length." Dissertation, North Texas University, Ann Arbor, MI.

*Dixon, Jo. 1995. "The Organizational Context of Criminal Sentencing." *American Journal of Sociology* 100(5):1157-98.

*Economic and Demographic Research. 1992. *An Empirical Examination of the Application of Florida's Habitual Offender Statute*. Tallahassee, FL: Joint Legislative Management Committee of the Florida Legislature.

*Engen, Rodney L. and Randy R. Gainey. 2000. "Modeling the Effects of Legally Relevant and Extralegal Factors Under Sentencing Guidelines: The Rules Have Changed." *Criminology* 38(4):1207-25.

*Engen, Rodney L., Randy R. Gainey, and Sara Steen. 1999. *Racial and Ethnic Disparities in Sentencing Outcomes for Drug Offenders in Washington State: FY1996 to FY1999*. Olympia, WA: Washington State Minority and Justice Commission.

Engen, Rodney L. and Sara Steen. 2000. "The Power to Punish: Discretion and Sentencing Reform in the War on Drugs." *American Journal of Sociology* 105(5):1357-95.

*Everett, Ronald S. and Barbara C. Nienstedt. 1999. "Race, Remorse, and Sentence Reduction: Is Saying You're Sorry Enough." *Justice Quarterly* 16(1):99-122.

Fears, Darryl. 2003. "U.S. Racial Labels Confuse Many Latino Immigrants." *The Seattle Times* January 18. Retrieved January 18, 2003 (http://seattletimes.nwsource.com/html/nationworld/134617859_latinorace18.html)

*Feeley, Malcolm M. 1979. *The Process Is the Punishment: Handling Cases in a Lower Criminal Court*. New York, NY: Russell Sage Foundation.

*Feimer, S., F. Pommersheim, and S. Wise. 1990. "Marking Time: Does Race Make a Difference? A Study of Disparate Sentencing in South Dakota." *Journal of Crime and Justice* 13 (1):86-102.

*Ferguson, Florence S. 1996. "Race, Urbanism, and Court Bureaucratization: An Empirical Examination of Conflict-Weberian Theories." Dissertation, Michigan State University, Ann Arbor, MI.

*Florida Sentencing Commission. 1997. *Sentencing Guidelines 1995-1996 Annual Report*. Tallahassee, FL: Florida Department of Corrections.

Glass, Gene V. 1976. "Primary, Secondary, and Meta-Analysis." *Educational Researcher* 5:3-8.

*Gorton, Joe and John L. Boies. 1999. "Sentencing Guidelines and Racial Disparity Across Time: Pennsylvania Prison Sentences in 1977, 1983, 1992 and 1993." *Social Science Quarterly* 80(1):37-54.

*Green, Edward. 1961. *Judicial Attitudes in Sentencing*. London: MacMillan.

Greenwald, Anthony G. 1975. "Consequences of Prejudice Against the Null Hypothesis." *Psychological Bulletin* 82(1):1-20.

Hagan, John. 1974. "Extra-Legal Attributes and Criminal Sentencing: An Assessment of a Sociological Viewpoint." *Law & Society Review* 8:357-83.

*Hagan, John and Ilene Bernstein. 1979. "Conflict in Context: The Sanctioning of Draft Resisters." *Social Problems* 27:109-22.

Hagan, John and Kristin Bumiller. 1983. "Making Sense of Sentencing: A Review and Critique of Sentencing Research." Pp. 1-54 in *Research on Sentencing: The Search for Reform*, vol. 2, Editors Alfred Blumstein, Jacqueline Cohen, Susan E. Martin, and Michael H. Tonry. Washington, D.C.: National Academy Press.

*Hagan, John, Ilene Nagel, and Celesta Albonetti. 1980. "The Differential Sentencing of White Collar Offenders in Ten Federal District Courts." *American Sociological Review* 45(5):802-20.

Hagan, John and Patricia Parker. 1985. "White-Collar Crime and Punishment: The Class Structure and Legal Sanctioning of Securities Violations." *American Sociological Review* 50(3):302-16.

*Hall, Edwin L. and Albert A. Simkus. 1975. "Inequality in the Types of Sentences Received by Native Americans and Whites." *Criminology* 13(3):199-222.

*Hanke, Penelope J. 1995. "Sentencing Disparities by Race of Offender and Victim: Women Homicide Offenders in Alabama, 1929-1985." *Sociological Spectrum* 15(3):277-97.

Harig, Thomas J. 1990. " The Influence of Extra-Legal Factors in Youthful Offender Adjudications." Dissertation, State University of New York at Albany, Ann Arbor, MI.

Harrison, Paige M. and Allen J. Beck. 2002. *Prisoners in 2001*. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Hasselblad, Vic and Larry V. Hedges. 1995. "Meta-Analysis of Screening and Diagnostic Tests." *Psychological Bulletin* 117(1):167-78.

Hedges, Larry V. 1982. " Estimation of Effect Size From a Series of Independent Experiments." *Psychological Bulletin* 92(2):490-499.

———. 1994. "Fixed Effects Models." Pp. 285-99 in *The Handbook of Research Synthesis*, Editors Harris Cooper and Larry V. Hedges. New York, NY: Russell Sage Foundation.

*Holmes, Malcolm D. and Howard C. Daudistel. 1984. "Race and Justice in the Southwest: The Sentencing of Anglo, Black, and Hispanic Defendants." *Social Science Quarterly* 65:265-77.

*Holmes, Malcolm D., Harmon M. Hosch, Howard C. Daudistel, Dolores A. Perez, and Joseph B. Graves. 1996. "Ethnicity, Legal Resources, and Felony Disposition in Two Southwestern Jurisdictions." *Justice Quarterly* 13(1):11-30.

Holmes, Malcolm D., Harmon M. Hosch, Howard C. Daudistel, Dolores A. Perez, and Joseph B. Graves. 1993. "Judges' Ethnicity and Minority Sentencing: Evidence Concerning Hispanics." *Social Science Quarterly* 74(3):496-506.

*Jankovic, Ivan. 1978. "Social Class and Criminal Sentencing." *Crime and Social Justice* 10(1):9-16.

*Johnson, Brian D. 2002. "Contextual Disparities in Sentencing: A Hierarchical Analysis of Sentencing Severity Across Pennsylvania Courts." Presentation at the University of Maryland, December 19, 2002.

*Kempf, Kimberly L. and Roy L. Austin. 1986. "Older and More Recent Evidence on Racial Discrimination in Sentencing." *Journal of Quanitative Criminology* 2(1):29-48.

Kleck, Gary. 1981. "Racial Discrimination in Criminal Sentencing: A Critical Evaluation of the Evidence with Additional Evidence on the Death Penalty." *American Sociological Review* 46(6):783-805.

*Klein, Stephen, Joan Petersilia, and Susan Turner. 1988. *Racial Equity in Sentencing*. Santa Monica, CA: RAND.

———. 1990. "Race and Imprisonment Decisions in California." *Science* 247:812-16.

Klepper, Steven, Daniel Nagin, and Luke-Jon Tierney. 1983. "Discrimination in the Criminal Justice System: A Critical Appraisal of the Literature." Pp. 55-128 in *Research on Sentencing: A Search for Reform*, vol. II, Editors Alfred Blumstein, Jacqueline Cohen, Susan E. Martin, and Michael H. Tonry. Washington, D.C.: National Academy Press.

*Kramer, John and Darrell Steffensmeier. 1993. "Race and Imprisonment Decisions." *Sociological Quarterly* 34(2):357-76.

*Kramer, John H. and Jeffery T. Ulmer. 1996. "Sentencing Disparity and Departures From Guidelines." *Justice Quarterly* 138(1):81-106.

*Kruttschnitt, Candace. 1980-1981. "Social Status and Sentences of Female Offenders." *Law & Society Review* 15(2):247-65.

*LaBeff, Emily E. 1978. "Sentence Length and Sentencing Variation: An Analysis and Clarification of a Problem in the Sociology of Law." Dissertation, Oklahoma State University, Ann Arbor, MI.

LaFree, Gary D. 1985. "Official Reactions to Hispanic Defendants in the Southwest." *Journal of Research in Crime and Delinquency* 22(3):213-37.

*Langan, Patrick A. 1999. "Sentence Reductions for Drug Traffickers for Assisting Federal Prosecutors." Unpublished manuscript.

Lipsey, Mark W. and David B. Wilson. 1993. "The Efficacy of Psychological, Educational, and Behavioral Treatment: Confirmation From Meta-Analysis." *American Psychologist* 48(12):1181-209.

———. 2001. *Practical Meta-Analysis*. Thousand Oaks, CA: Sage Publications.

Long, J. S. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage.

Lotz, Roy and John D. Hewitt. 1977. "The Influence of Legally Irrelevant Factors on Felony Sentencing." *Sociological Inquiry* 47(1):39-48.

Mann, Charles C. 1994. "Can Meta-Analysis Make Policy?" *Science* 266(5187):960-962.

*Martin, Marcus E. and Melissa Stimpson. 1997-1998. "Women, Race, and Sentencing in Oklahoma: A Preliminary Analysis." *Journal of the Oklahoma Criminal Justice Research Consortium* 4(August).

*McDonald, Douglas C. and Kenneth E. Carlson. 1993. *Sentencing in the Federal Courts: Does Race Matter? The Transition to Sentencing Guidelines, 1986-1990*. Washington, DC: U.S. Department of Justice, Office of Justice Programs and Bureau of Justice Statistics.

*Meyer, Jon'a and Tara Gray. 1997. "Drunk Drivers in the Courts: Legal and Extra-Legal Factors Affecting Pleas and Sentences." *Journal of Criminal Justice* 25(2):155-63.

Miethe, Terance D. and Charles A. Moore. 1987. *Evaluation of Minnesota's Felony Sentencing Guidelines: Final Report*.

Moore, Charles A. and Terence D. Miethe. 1986. "Regulated and Unregulated Sentencing Decisions: An Analysis of First-Year Practices Under Minnesota's Felony Sentencing Guidelines." *Law & Society Review* 20(2):253-77.

Musto, David F. 1987. *The American Disease: Origins of Narcotic Control*. New York: Oxford University Press.

*Myers, Laura B. 1990. "A Comparison of Legal and Extra-Legal Factors Affecting Criminal Court Decision-Making." Dissertation, Florida State University, Ann Arbor, MI.

*Myers, Martha A. 1979. "Offended Parties and Official Reactions: Victims and the Sentencing of Criminal Defendants." *Sociological Quarterly* 20:529-40.

———. 1987. "Economic Inequality and Discrimination in Sentencing." *Social Forces* 65(3):746-66.

*———. 1989. "Symbolic Policy and the Sentencing of Drug Offenders." *Law & Society Review* 23(2):295-315.

Myers, Martha A. and Susette M. Talarico. 1986. "The Social Contexts of Racial Discrimination in Sentencing." *Social Problems* 33(3):236-51.

*———. 1987. *The Social Contexts of Criminal Sentencing*. New York: Springer-Verlag.

*Nagel, Stuart. 1969. *The Legal Process from a Behavioral Perspective*. Homewood, Ill.: Dorsey Press.

Nagin, Daniel S. and Raymond Paternoster. 1991. "On the Relationship of Past to Future Participation in Delinquency." *Criminology* 29(2):163-89.

Nelson, James F. 1989. "An Operational Definition of Prior Criminal Record." *Journal of Quantitative Criminology* 5(4):333-52.

*———. 1992. " Hidden Disparities in Case Processing: New York State, 1985-1986." *Journal of Criminal Justice* 20(3):145-60.

———. *Disparities in Processing Felony Arrests in New York State*, 1990-1992. Albany, NY: New York State Division of Criminal Justice Services.

Overton, Randall C. 1998. "A Comparison of Fixed-Effects and Mixed (Random-Effects) Models for Meta-Analysis of Moderator Variable Effects." *Psychological Methods* 3(3):354-79.

*Pasko, Lisa. 2002. "Villain or Victim: Regional Variation and Ethnic Disparity in Federal Drug Offense Sentencing." *Criminal Justice Policy Review* 13(4):307-28.

*Petersilia, Joan. 1983. *Racial Disparities in the Criminal Justice System*. Santa Monica, CA: RAND.

*Peterson, Ruth D. 1988. "Youthful Offender Designations and Sentencing in the New York Criminal Courts." *Social Problems* 35(2):111-30.

*Peterson, Ruth D. and John Hagan. 1984. "Changing Conceptions of Race: Towards an Account of Anomalous Findings and Sentencing Research." *American Sociological Review* 49(1):56-70.

*Pope, Carl E. 1975. *Sentencing of California Felony Offenders*. Washington, D.C.: U.S. Department of Justice.

*Pruitt, Charles R. and James Q. Wilson. 1983. "A Longitudinal Study of the Effect of Race and Sentencing." *Law & Society Review* 17(4):613-35.

Raudenbush, Stephen W. 1994. "Random Effects Models." Pp. 302-21 in *The Handbook of Research Synthesis*, Editors Harris Cooper and Larry V. Hedges. New York, NY: Russell Sage Foundation.

*Rodriguez, Nancy. 1998. "Racial Disparity, Patterned Offenses, and Practical Impacts: An Assessment of the Persistent Offender Law in the State of Washington." Dissertation, Washington State University, Ann Arbor, MI.

*Sealey, Beverly. 1995. " Disparity in Sentencing: The Influence of Race on Sentencing Outcomes Among Black and White Males Incarcerated for Felony Offenses of Robbery in Massachusetts." Dissertation, Brandeis University, Ann Arbor, MI.

Sellin, Thorsten. 1928. "The Negro Criminal: A Statistical Note." *The Annals of the American Academy of Political and Social Science* 140:52-64.

*Simon, Lenore M. J. 1996. "The Effect of the Victim-Offender Relationship on the Sentence Length of Violent Offenders." *Journal of Crime and Justice* 19(1):129-48.

Smith, Mary L. 1980. "Publication Bias and Meta-Analysis." *Evaluation in Education* 4(1):22-24.

*Souryal, Claire and Charles Wellford. 1999. "Sentencing Disparity and Sentencing Guidelines." Pp. 199-230 in *The Criminology of Criminal Law: Advances in Criminological Theory*, vol. 8, Editors William S. Laufer and Freda Adler. New Brunswick, NJ: Transaction Publishers.

Spohn, Cassia. 2000a. "Thirty Years of Sentencing Reform: The Quest for a Racially Neutral Sentencing Process." Pp. 427-501 in *Policies, Processes, and Decisions of the Criminal Justice System: Criminal Justice 2000*, vol. 3, Editor Julie Horney. Washington, D.C.: National Institute of Justice.

*———. 2000b. *Offender Race and Case Outcomes: Do Crime Seriousness and Strength of Evidence Matter? (Final Report to NIJ)*. Washington, D.C.

*Spohn, Cassia and Jerry Cederblom. 1991. "Race and Disparities in Sentencing: A Test of the Liberation Hypothesis." *Justice Quarterly* 8(3):305-27.

*Spohn, Cassia and Miriam DeLone. 2000. "When Does Race Matter?: An Analysis of the Conditions Under Which Race Affects Sentence Severity." *Sociology of Crime, Law, and Deviance* 23-37.

*Spohn, Cassia, Miriam DeLone, and Jeffrey Spears. 1998. "Race/Ethnicity, Gender, and Sentence Severity in Dade County: An Examination of the Decision to Withhold Adjudication." *Journal of Crime and Justice* 21(2):111-38.

*Spohn, Cassia, John Gruhl, and Susan Welch. 1981-1982. "The Effect of Race on Sentencing: A Re-Examination of an Unsettled Question." *Law & Society Review* 16(1):71-88.

*Spohn, Cassia and David Holleran. 2000. "The Imprisonment Penalty Paid by Young, Unemployed Black and Hispanic Male Offenders." *Criminology* 38(1):281-306.

*Steffensmeier, Darrell and Stephen Demuth. 2000. "Ethnicity and Sentencing Outcomes in U.S. Federal Courts: Who Is Punished More Harshly?" *American Sociological Review* 65(5):705-29.

Steffensmeier, Darrell, Jeffery Ulmer, and John Kramer. 1998. "The Interaction of Race, Gender, and Age in Criminal Sentencing: The Punishment Cost of Being Young, Black and Male." *Criminology* 36(4):763-97.

Steffensmeier, Darrell J. 1980. "Assessing the Impact of the Women's Movement on Sex-Based Differences in the Handling of Adult Criminal Defendants." *Crime and Delinquency* 26(3):344-57.

Tonry, Michael. 1995. *Malign Neglect: Race, Crime, and Punishment in America*. New York: Oxford University Press.

———. 1996. *Sentencing Matters*. New York, NY: Oxford University Press.

*Truitt, Linda. 1997. "Analysis of Racial Disparity in Felony Criminal Court Sentencing." Dissertation, University of California Irvine, Ann Arbor, MI.

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

*Ulmer, Jeffery T. and John H. Kramer. 1996. "Court Communities Under Sentencing Guidelines: Dilemmas and Formal Rationality and Sentencing Disparity." *Criminology* 34(3):383-408.

*Unnever, James D. 1982. "Direct and Organizational Discrimination in the Sentencing of Drug Offenders." *Social Problems* 30:212-25.

*Unnever, James D., Charles E. Frazier, and John C. Henretta. 1980. "Race Differences in Criminal Sentencing." *The Sociological Quarterly* 21:197-206.

*Walsh, Anthony. 1991. "Race and Discretionary Sentencing: An Analysis of ‘Obvious’ and ‘Non-obvious’ Cases." *International Journal of Offender Therapy and Comparative Criminology* 31(1):7-19.

*Walters, Michael J. 1982. "Sequential Decision-Making in the Criminal Court: The Influence of Prior Court Outcomes on Sanctioning and Their Relationship to Discrimination." Dissertation, University of Cincinnati, Ann Arbor, MI.

Wang, Morgan C. and Brad J. Bushman. 1999. *Integrating Results Through Meta-Analytic Review Using SAS Software*. Cary, NC: SAS Institute, Inc.

Welch, Susan, Cassia Spohn, and John Gruhl. 1984. "Dismissal, Conviction, and Incarceration of Hispanic Defendants: A Comparison with Anglos and Blacks." *Social Science Quarterly* 65:257-64.

*———. 1985. "Convicting and Sentencing Differences Among Black, Hispanic, and White Males in Six Localities." *Justice Quarterly* 2(1):67-80.

Wilbanks, William. 1987. *The Myth of a Racist Criminal Justice System*. Monterey, CA: Brooks/Cole.

*Wonders, Nancy A. 1990. "Courts in Context: A Sociological Approach to Sentencing Disparity." Dissertation, State University of New Jersey, New Brunswick, Ann Arbor, MI.

*Wooldredge, John D. 1998. "Analytical Rigor in Studies of Disparities in Criminal Case Processing ." *Journal of Quantitative Criminology* 14(2):155-79.

*Wooldredge, John D., Fritz Rauschenberg, Timothy Griffin, and Travis Pratt. 2002. *The Impact of Ohio's Senate Bill 2 on Sentencing Disparities*. Cincinnati, OH: Center for Criminal Justice Research, University of Cincinnati.

*Zatz, Marjorie S. 1984 "Race, Ethnicity, and Determinate Sentencing: A New Dimension to an Old Controversy." *Criminology* 22(2):147-71.

———. 1987. "The Changing Forms of Racial/Ethnic Biases in Sentencing." *Journal of Research in Crime and Delinquency* 24(1):69-92.

Zatz, Marjorie S. and John Hagan. 1985. "Crime, Time, and Punishment: An Exploration of Selection Bias in Sentencing Research." *Journal of Quantitative Criminology* 1(1):103-26.

*Zimmerman, Sherwood and Bruce Frederick. 1984. "Discrimination and the Decision to Incarcerate." Pp. 315-34 in *The Criminal Justice System and Blacks*, Editor Daniel Georges-Abeyie. New York: Clark Boardman.