JOURNAL OF ETHNICITY IN CRIMINAL JUSTICE
2016, VOL. 0, NO. 0, 1–21
http://dx.doi.org/10.1080/15377938.2016.1187239

Routledge
Taylor & Francis Group

It's not all black and white: A propensity score matched, multilevel examination of racial drug sentencing disparities

Richard J. Stringer[a] and Melanie M. Holland[b]

[a]Department of Sociology and Criminal Justice, Old Dominion University, Norfolk, Virginia, USA;
[b]Department of Criminal Justice, Pennsylvania State University, Altoona, Pennsylvania, USA

ABSTRACT

This study aims to alleviate some of the mixed findings throughout the literature on racial disparities in sentencing outcomes by utilizing propensity score matching and multilevel modeling to assess racial drug sentencing disparities in state courts from 2000–2012. The findings illustrate the effect of race on sentencing varies significantly across states, and aggregate factors impact this relationship. Specifically, although differential offending, minority population, and arrests do not alleviate disparities, they are moderators that explain variance across states. Finally, aggregate socioeconomic factors such as poverty and education are also significant moderators that indicate the importance of structural disadvantage in sentencing outcomes.

ARTICLE HISTORY
Received 2 October 2015
Revised 20 February 2016
Accepted 2 March 2016

KEYWORDS
Drugs; race; courts; sentencing

Introduction

The criminal justice system provides a means of holding individuals liable for their wrongdoings. An effective system can incapacitate dangerous offenders, reduce future victimization, and promote public perceptions of safety. A biased one, however, generates concerns of the legitimacy of the institution. Justice is not achieved when the courts demonstrate prejudices. Currently, African-Americans are disproportionately arrested, charged, imprisoned, granted longer sentences, and represented in capital offenses (Everett & Wojtkiewicz, 2002; Mitchell, 2005). As prison populations soar, the number of African-American men entering this system has become astounding; African-Americans represent only 13.2% of the population (United States Census Bureau, 2013), but 28.3% of all arrests and 30.4% of drug abuse violations (Federal Bureau of Investigation, 2013c). African-American men who are under the age of 40 years old demonstrate an incarceration rate of 11.5% (Western, 2006). Such inequalities in sentencing can be viewed as an expansion and continuation of other historical injustices faced by disadvantaged people of color (e.g., Alexander, 2010; Western, 2006). While research has examined the

CONTACT Richard J. Stringer rstringe@odu.edu Department of Sociology and Criminal Justice, Old Dominion University, Norfolk, VA 23529, USA.

effects of race on sentencing outcomes to some extent, few recent studies have assessed the effect of region and state level factors on this relationship.

African-Americans have suffered historically from subjugation, repression, and marginalization throughout the United States; however, nowhere has this racialized exploitation been as evident as it has been in the South. The Southern history of slavery and Jim Crow are two notorious examples that are demonstrative of the history of race-based subjugation within this region (Alexander, 2010). Though extremely limited and dated, contemporary research examining prejudicial attitudes across region suggest that racist attitudes continue to linger within the South. According to Taylor (1998), higher populations of African-Americans are related to more prejudicial attitudes in regions outside of the South, with Southern states demonstrating much less variability in prejudices. Therefore, while local factors are influential outside of the South, Southern prejudices appear more inflexible. Although other state level predictors are important, Southern states continue to demonstrate unique sentencing practices compared to other regions. In a study examining the effects of several state level predictors on sentencing severity, Michalowski and Pearson (1990) found that belonging to a Southern state was the strongest state level predictor.

While the abolition of slavery and Jim Crow practices have greatly reduced blatant racism in the South, many argue that race-implicit terminology and symbolism have subverted overt discrimination. Particularly, racial minority has become synonymous with criminality, likely influenced by the overrepresentation of African-American felons which Michelle Alexander (2010) equates to "The New Jim Crow." Like the Jim Crow policies prior to the Civil Rights Movement and slavery before then, a cultural lag appears within the South when ensuring the fair treatment of African-Americans (Michalowski & Pearson, 1990; Taylor 1998), suggesting that regional variation may be of particular importance when examining racial inequalities within the United States. This study contributes to the existing literature by exploring state level predictors of racial disparities in sentencing. The examination of state level and region specific variations in sentencing outcomes across race not only presents the opportunity to better understand the link between structural level factors and micro-level decision making, but by identifying population triggers of inequality, these findings can point policy makers in the direction of locations most in need of policy reform.

Although mass incarceration has become pervasive throughout the United States, the growing number of individuals thrust into this system may be partially driven by the Southern region. According to the 2013 UCR, the Southern region reported the greatest number of arrests (3,451,627) and leads the regions with the highest arrest rates (4,096.9). The remaining regions demonstrate an arrest rate ranging from 3,084.1 (Northeast) to 3,647.2 (Midwest), with the West falling in between (3,618.6). The South also leads the arrest rates in murders and larcenies (Federal Bureau of Investigation, 2013a). Of particular interest, however, is the finding that the South demonstrates the highest rate of drug abuse violations

(540.7) compared to the Northeast (446.3), Midwest (463.3), and West (483.4) (Federal Bureau of Investigation, 2013a). Of drug arrests in 2013, the Southern region reported the highest number of synthetic and manufactured drug sales and possessions (Federal Bureau of Investigation, 2013b). As a result of these copious arrest statistics, this study will focus on drug arrests.

Although previous research has examined the effects of race on sentencing outcomes, several important gaps in the literature remain. For instance, it remains unclear whether racial disparities are a product of differential offending. In order to effectively compare African-American and Caucasian offenders, this study will match the two groups using a quasi-experimental approach. Furthermore, research has yet to adequately examine state level and regional variations in sentencing outcomes. This study fills this gap by including contextual variables which can aid policy makers in their attempt to limit the manifestation of racial discrimination in the courtroom. Using longitudinal data collected by the Bureau of Justice Statistics' National Corrections Reporting Program from 2000–2012, this study contributes to the current literature by examining whether African-American drug offenders receive more severe sentences than their matched Caucasian counterparts while controlling for legally relevant variables, and whether such variability is dependent upon the U.S. state/region in which the case presides.

Literature review

Causal explanations for racial disparities

Research examining the relationship between racial discrimination within the criminal justice system has been identified as one of the dominant topics within sentencing literature according to Zatz (1987), and this continues to be true today. Of research examining differential sentencing outcomes across races, research consistently acknowledges that African-Americans are disproportionately represented in the penal system (Pratt, 1998). In fact, of decision making stages, Stolzenberg, D'Alessio, & Eitle (2013) found that racial disparities are only significant for incarceration and sentence length decisions in contrast to "granting financial release, denying bail, bail amount, being held on bail, pretrial incarceration, and whether the defendant was adjudicated as a felon" (p. 286). Therefore, research examining racial inequality in the criminal justice system may best assist the existing literature by focusing on incarceration-related sentencing outcomes. Though the purpose of this study is not to test specific theoretical explanations, but rather to examine case and state level predictors of racial inequality, it is worth noting that there are several causal explanations for racial inequality within the sentencing literature which can be divided into one of three arguments.

The first causal explanation for the racial disparity in incarceration is that African-Americans disproportionately commit more severe crimes. A great deal of empirical research has found support for this differential offending thesis (Barnes & Kingsnorth, 1996; Blumstein, 1982; Crutchfield, Fernandez, & Martinez, 2010;

Harris, Steffensmeier, Ulmer, & Painter-Davis, 2009; Kleck, 1981; Pratt, 1998; Sorensen, Hope, & Stemen, 2003). One of the most frequently cited studies on the causes of racial disparities within prisons supports the differential offending argument, and finds that 80% of racial disproportionality is a product of differential criminality (Blumstein, 1982).

The second explanation for the race gap is often referred to as the direct impact argument. As the antithesis of the differential offending rationalization, numerous studies have found that racial inequality continues to exist even when controlling for legally relevant variables, suggesting that racial inequality may be a product of residual prejudices (e.g., Arvanites, 2014; Bourassa & Andreescu, 2009; Hawkins & Hardy, 2015; Mitchell, 2005; Weitzer, 1996). In a study examining sentencing outcomes in a county in the Southern state of Kentucky, Bourassa & Andreescu (2009) argue that only one-third of racial disparities can be explained by case characteristics. Findings of discrimination do not appear to be limited to this one state, however. Mitchell's (2005) meta-analysis examined 71 studies dating from 1929 to 2000. His study found that African-Americans generally received harsher sentences than their Caucasian counterparts. This sentencing disparity continued to be constant even when controlling for offense severity and criminal history (Mitchell, 2005). According to this study, findings claiming that disproportionate sentencing is a result of disproportionate offending are erroneous, suggesting that racial discrimination may, in fact, exist at the sentencing level.

Yet a third and final delineation of the cause of racial disparities in incarceration suggests that extraneous, potentially race-neutral, factors are to blame for the disproportionality, rejecting both of the former extremes. These interactions may occur at the individual level with some theorists arguing racial disparities in prison are actually a product of differential attributions of blameworthiness, dangerousness, and practical constraints as suggested in Focal Concern Theory (Steffensmeier et al., 1998). Steffensmeier et al. (1998) argue that stereotypes of African-Americans as criminogenic and dangerous may explain the harsher sentencing outcomes. The interaction effect may also occur at the structural level. In Schlesinger's (2011) study examining male state prison admission records, the author found that mandatory imprisonment policies increased admissions of African-American men more so than their Caucasian counterparts. The demography of a community might condition the effects of race on sentencing outcomes (Arvanites & Asher, 1995; Sorenson et al., 2003). Others have argued that urban and rural environments differentially impact the importance of race (Mitchell, 2005).

The incongruity of explanations explaining racial disparities within sentencing outcome may actually be artifactually generated as a result of differential methodologies. Marjorie Zatz's (1987) comparison of the preexisting literature on racial disparities in sentencing outcomes across four waves from 1930–1980 finds that the manifestation of racial discrimination has changed across time according to variations in analytical strategies. Studies conducted from 1930 through the mid-1960s (wave 1) demonstrate "clear and consistent bias against nonwhite in

sentencing," (Zatz, 1987, p. 82) while wave 2 (1960s through the 1970s) suggested that no discrimination was present. The third wave examined included studies published during the 1970s and 1980s (wave 3) that often suggested that African-Americans receive more severe outcomes when controlling for important variables, and the last wave of data in this study, most of which was published in the 1980s (wave 4), took into account determinate sentencing and shifting discretion spreading throughout this decade, suggesting that racial discrimination may have become a product of prosecutors as gatekeepers (Zatz, 1987).

In order to effectively identify whether racial disparities in sentencing outcomes are a result of differential offending, overt discrimination, or race-neutral interactions, research must adequately control for legal and extralegal factors across race. Propensity score matching is a promising technique to achieve this because it can match groups according to various criteria. Only one known study has applied propensity score matching to the analysis of racial disparities in sentencing outcomes. Jordan & Freiburger (2014) find that race effects continue to exist between African-American and Caucasian offenders for jail sentences, in relation to prison and probation, but do not differ according to prison length. Though an important contribution to the literature, the State Court Processing data analyzed in this study was collected using samples within large urban counties, limiting the generalizability of the findings. Furthermore, the study did not examine state level predictors. Using propensity score matching, this study uses quasi-experimental designs to more thoroughly control for between race differences as well as multilevel modeling to examine state level variations using the National Corrections Reporting Program which includes statewide statistics across the United States.

Race effects on drug offenses

While racial disparities appear to be an ever present reality throughout the criminal justice system and across a variety of offenses, racial disadvantages may be more pronounced for drug offenses. The incarceration boom, culminating in the 1980s with the largest imprisonment increase, has been somewhat driven by drug arrests with 33% of this increase due to drug offenders (Schoenfeld, 2012). The incarceration increase for drug offenses resulted in an increase of African-American imprisonment that is 4 times that of Caucasians (Arvanites, 2014). While the state prison population growth appears to be in the process of tapering off, drug arrests continue to disproportionately affect the African-American community (Alexander, 2010; Massey, 2007).

Arvanites (2014) claims that it is unlikely that the disproportionate increase in African-American drug arrests from 1980–2005 is a result of differential drug use. For non-Federal courts, racial disparities appear the most pronounced for drug offenses (Mitchell, 2005). While legal factors fail to explain all of the racial variations in sentencing outcomes (roughly 80%), such variables explain even less variance for drug offenses (less than 50%) (Crutchfield et al., 2010). Surprisingly, the inclusion of mandatory drug sentencing fails to reduce racial disparities.

Mandatory drug sentencing increases drug admissions for both African-American and Caucasian offenders. These laws, however, disproportionately increase African-American incarceration. The effects of race on incarceration are not as consistent, however, as they are on other mandatory policies (Schlesinger, 2011). Therefore, this study will specifically examine drug offenders.

Effects of race across place

Although extensive research has addressed racial inequality within sentencing decisions, few studies have investigated regional and/or geographic variations of racial sentencing disparities despite the historical propensity of the South to demonstrate and support policies and attitudes that appear particularly prejudicial. It is probable that racial disparities vary greatly according to region as a result of differential sociohistorical context; therefore, studying various jurisdictions is essential. Of the few studies accounting for meso and macro level contexts, unique distinctions appear to emerge across place. For instance, several authors have noted that racial profiling (Lyons et al., 2013) and sentencing disparities (Michell, 2005) are more highly emphasized in rural areas compared to urban ones. Increases in poverty also appear to be related sentencing outcomes (Beckett & Western, 2001). Arvanites and Asher (1995) found that the percent of non-Caucasians per state is directly related to imprisonment rates, though the importance of racial demography varied between the South and the rest of the United States. According to Beckett & Western (2001), the effects of the percentage of minorities is modest, although it does appear to be growing. Sorensen et al. (2003) argue that the higher rates of African-American crime in locations with higher concentration of African-Americans is a product of differential involvement.

According to Pasko (2002), varying legislation, as well as extralegal biases unique to specific regions, offers Caucasian men and women with an advantage in the courtroom. These differences are often applied through judicial discretion, suggesting that guidelines are applied differently across regions. The study further found that sentence length was closely tied to differing regional attitudes (Pasko, 2002). Using 2002 prison admission data for drug offenses, Arvanites (2014) found that African-Americans have higher incarceration rates for drug offenses in counties with less segregation, potentially a product of less visibility. The differing attitudes Pasko (2002) references appear to demonstrate a hardening effect within the South. According to Gary Kleck's 1981 study, the South was the only region within the Unites States in which African-Americans guilty of homicide are more likely to receive the death penalty than their Caucasian counterparts. While other state-level predictors, including state revenue, are important predictors of severity, it is the Southern variable that is the strongest predictor of sentencing severity in Michalowski & Pearson's (1990) study examining data from 1970 and 1980. In a more recent study of punishment severity and region, Pritchard & Wiatrowski (2008) found that Southern states are more likely to have and implement the death penalty. Unfortunately, of these few

studies that have incorporated contextual level data within sentencing severity data, most have utilized decades old data and often failed to include region. In order to assess sentencing disparities across place, this study includes several state level predictors as well as region to examine sentence length using longitudinal data from 2000–2012.

Methodology

Drawing from previous research examining racial disparities in sentencing decisions, this study identifies two research questions.

R1: Do African Americans receive disproportionately longer sentences for drug offenses, while controlling for legally relevant variables, than Caucasian drug offenders?
R2: Does this relationship vary by state and/or region of the country?

Because court decisions do not occur within a sociological vacuum, individual sentencing outcomes are expected to be influenced by the local zeitgeist of the location in which the decision occurs, suggesting the likelihood of potential clustering effects. Therefore, multilevel modeling will be used to examine cases nested within states.

Level I data source

This study utilized secondary data from the National Corrections Reporting Program (NCRP) for the first level of analysis (U.S. Department of Justice, 2014). The NCRP data are compiled from reported state records for prison admissions, releases, and custody within the reporting period of 2000 through 2012. Cases were excluded from analysis that did not indicate that a drug offense was the primary offense which resulted in a total of 2,177,092 cases. A random sample of about 10% of the population of those incarcerated was taken for analysis herein. Unfortunately, seven states have either not reported, or inconsistently reported, on drug sentences within this term and are therefore not included in the term file (ICPSR, 2014); thus Arkansas, Arizona, Connecticut, Hawaii, New Mexico, Virginia, and Vermont are not represented in these data. Propensity score matching is therefore utilized in order to correct for sampling bias within the data and increase internal validity of the findings.

Due to missing data on a few variables, multiple imputation methods were utilized in order to replace these missing values and allow for further control of the statistical analysis (Allison, 2001). Descriptive statistics on these data are presented in Table 2. The variables with imputed values are the number of counts for the offense, community corrections supervision when the current offense was committed, prior felony convictions, education, and determinate sentencing. An exploratory analysis of the number of counts revealed that offenders with more than two counts were outliers, and thus this variable was recoded as a dummy variable for multiple counts prior to imputation. Community corrections supervision, a prior

felony conviction, and determinate sentencing were all dichotomous variables originally. Furthermore, education was constructed as an ordinal variable ranging from 1–5 that represents no high school education, some high school, high school graduate, some college, and college graduate.

Several indicators were used to predict the missing values for each variable. They include offense type, gender, age, prior felony convictions, education, sentence length, total prior incarceration time, the number of counts for the current offense, and community corrections supervision. Finally, the imputation models were constrained to return an integer within the range of the original data due to the categorical and ordinal nature of these data. Descriptive statistics on the final dataset with the missing values imputed are presented in Table 2.

Level II data sources

Level II data represented several state level factors hypothesized to effect sentencing decisions. Descriptive statistics for these data are presented in Table 1. These data were procured from the 2010 census of the United States that represent the demographic and economic composition of each state (Census, 2010). Indicators have been included for the total state population, the percentage of the population that is African-American, Caucasian, do not have a high school degree, and the percent below the poverty level. Rather than operationalize Southern states as those that fought for the confederacy in the civil war, a more recent definition of Southern state was used. This study used the 17 states that engaged in state-sponsored racial segregation. Since much has changed since the civil war, this is a better measure of sentiments toward the "Southern way of life" (Katznelson, 2005, p. 17). Katznelson (2005) argues that legal state-sponsored racial segregation was practiced in seventeen states until the mid-twentieth century. These states include Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.

Furthermore, this study utilized data that represent the percentage of persons who indicated that they had ever used marijuana within a state that are African-American between 2002 and 2011 (United States Department of Health and Human Services, 2013). Because marijuana is the most commonly used illicit drug within the

Table 1. Descriptive statistics of level II data.

	n	Range	Mean	Standard Deviation
Southern State	38	1	0.37	0.49
Population Percent Black	38	36.4	11.24	3.36
Determinate Sentencing State	38	1	0.34	0.48
Percent Below Poverty Level	38	12.5	14.93	2.8
Percent No High School Deg.	38	11.1	13.15	3.2
Marijuana Use Percent Black	38	7.6	1.99	2.3
Marijuana Arrests Percent Black	38	67.9	27.17	19.83

United States, represents about half of the drug arrests, is argued to be the bedrock of the drug war, and enforcement of marijuana laws have been associated with racial disparities, it was utilized herein as a proxy for the percent of drug users that are African-American within a state (see, e.g., Gerber, 2004; Golub, Johnson, & Dunlap, 2007; Musto, 1999; United States Department of Health and Human Services, 2011). This study also utilized UCR data to control for the percent of total persons arrested for marijuana that are African-American within a state (U.S. Department of Justice, 2011). Finally, a dichotomous variable was constructed using information in the state fact sheets in the codebook to distinguish between determinate sentencing and indeterminate sentencing states (see ICPSR, 2014, Appendix A).

Propensity score matching

Offender level cases utilized in level 1 were divided and matched into 2 quasi-experimental groups using propensity score matching techniques in STATA 12. The goal of this method is to have both a control and a treatment group that have been systematically matched on several criteria known to significantly predict treatment. This method allows for greater internal validity and the ability to make causal inferences by modeling counterfactual trends within the data and limiting the potential for spurious effects due to the systematic matching technique (Guo & Fraser, 2014). Thus, the goal was to create a control group made up of Caucasians and a treatment group of African-Americans. As such, a dummy variable was created that compared African-Americans to Caucasians as the reference category. Offenders who were identified as any other race or ethnicity were coded as missing during this process.

Because this study utilized multilevel post-matching analysis in order to analyze cases nested within states, the propensity score matching process was amended. Although across cluster matching can be conducted, this process does not account for heterogeneity across clusters, and has greater limitations compared to within-cluster matching (Steiner, Kim, & Thoemmes, 2013; Thoemmes & West, 2011). As such, within-cluster matching was conducted consistent with the procedures set forth by Thoemmes & West (2011). This was accomplished by dividing the cumulative drug offender file into several separate files by state, prior to matching. This allowed for matching to be conducted within each cluster or state in a manner consistent with the approach of Thoemmes & West (2011). This process reduces the potential for bias in the matching procedure by assuring that aggregate or cluster level factors are held constant when matching treatment and control cases within clusters (Kim & Steiner, 2015). Because propensity score matching requires a large sample size within each cluster in order to adequately perform matching, cases within four of the states could not be included due to low sample sizes. These states were Maine (470), Montana (700), New Hampshire (2,398), and Wyoming (1,773).

Next, a binary logistic regression was performed using each state file in order to identify statistically significant predictors of being in the treatment. This analysis

revealed several statistically significant predictors of being in the treatment category. These outcomes varied by state, however. The significant predictors of treatment that were utilized to conduct the matching were age at admission to prison, admission year, prior felony convictions, gender, the number of counts for the current offense, level of education, prior jail and prison time, and community corrections supervision prior to prison sentencing. Thus, these variables were used to match the two groups together using psmatch2 in Stata 12. The propensity score analysis was conducted as a logit analysis versus the default prohibit model. In order to obtain the best match possible a second matching analysis was conducted utilizing a caliper that ranged from 0.02 to 0.11 which was calculated using the standard method of SD*0.25 (with SD representing the standard deviation of the propensity scores from the first matching analysis) (Guo & Fraser, 2014). This method was repeated for all of the level 1 state files. All unmatched cases were discarded from the data files which resulted in a decrease in sample size (n = 144,263). Because this resulted in a significant amount of power that would be cause for concern when interpreting statistical significance, a random sample of approximately 10% of the total matched sample was drawn. Descriptive statistics on the level I data are presented in Table 2.

Hierarchical linear modeling

The study utilized multilevel hierarchical linear modeling in the HLM 7 program due to the scale makeup of the sentence length dependent variable. The sentence length variable was truncated at 470 months in a manner consistent with prior research and the idea that any sentence above this is a life sentence (see Ulmer & Light, 2011). Furthermore, both the sentence length and prior incarceration time variables were transformed using a logarithmic transformation (log+1) in order to

Table 2. Descriptive statistics of level I propensity score matched data.

	n	Range	Mean	Standard Deviation
Treated (African American)	144,263	1	0.60	0.49
Sentence Length in Months	144,263	470	62.87	83.76
Year of Admission	144,263	42	2004	5.1
Age at Admission	144,263	87	34.43	9.8
Male	144,263	1	0.87	0.341
Possession of Marijuana	144,263	1	0.02	0.131
Possession of Heroin	144,263	1	0	0.058
Possession of Cocaine/Crack	144,263	1	0.03	0.167
Possession of Other Drug	144,263	1	0.06	0.241
Trafficking Marijuana	144,263	1	0.03	0.181
Trafficking Heroin	144,263	1	0	0.038
Trafficking Cocaine\Crack	144,263	1	0.03	0.181
Trafficking Other Drug	144,263	1	0.11	0.314
Multiple Counts	144,263	1	0.10	0.3
Determinate Sentence	144,263	1	0.67	0.471
Prior Felony Conviction	144,263	1	0.36	0.481
Prior Incarceration in Months	144,263	15	5.6	3.93
Education	144,263	4	2.45	0.778

normalize them prior to analysis. Initially, a third level of analysis was attempted that represented the four regions of the country, Midwest, Northeast, South, and West as defined by the U.S. Census (Census, 2010a). However, the introduction of a third level of analysis resulted in an insignificant *p*-value, which indicates that most of the variance has likely been accounted for at level two, and a third level of analysis was not appropriate here. As such, the individual offender level files are used as the level I unit of analysis, while the state level file was added as level II.

First, a base or null model was run for each of the three datasets so that the variance explained in the subsequent models could be calculated. Next, fixed effects models were estimated. A statistically significant *p*-value in the variance components estimation illustrated in Tables 3 and 4 indicates that the results do significantly differ from the results that would be obtained from a single level fixed effects logistic regression; thus, HLM is the appropriate method of analysis for these data (Raudenbush & Bryk, 2002). Random effects models were estimated by opening the error term for each coefficient one at a time and then examining the statistical significance of the final estimation of variance components. Random effects that were statistically significant were allowed to vary, while others that were not statistically significant were set to fixed in the final model.

Findings

Tables 3 and 4 present the findings from the multivariate hierarchical linear models that predict sentence length. These models include all 38 good reporting states (see ICPSR, 2014, Appendix A), and 144,263 individual level cases at level 1. The reliability estimates for all seven multilevel models indicate that there is little sampling error and that all or 99% of the variance in sentence length is explainable by the statistical models. According to the null model the intraclass correlation coefficient is 0.397. Thus, approximately 40% of the variance in sentence length is at the state level, and the remaining 60% remains at the individual level. The variance explained between and within each state is indicated by the level I and II r-squared values. These values were calculated using a method that accounts for varying within cluster sample sizes by utilizing a formula that incorporates the harmonic mean (see Snijders & Bosker, 1994, 2004). The fixed effects level 1 models account for approximately 8% of the variance in sentence length. However, since chi-squared statistics were significant ($P < 0.001$) for all indicators when allowing them to vary across states, random effects models are also presented which account for about 26% of the variance in sentence length at level I. The variance explained at level II varied from 80 to 98% when the effect of level II variables was tested on the treatment effect of being African American.

The treatment effect of race (African American) was a significant predictor of sentence length in most of the fixed effect models; however, when this indicator was allowed to vary across states it was no longer statistically significant

Table 3. Hierarchical linear models predicting log sentence length in months.

	Model I Level I Fixed	Model II Level I Random	Model III Level II
Treated (African American)	0.052*** (0.006)	0.042 (0.026)	—
Age at Admission	0.002*** (0.000)	0.004** (0.001)	—
Male	0.167*** (0.008)	0.165*** (0.027)	—
Education	0.017*** (0.004)	0.019* (0.007)	—
Prior Felony Conviction	0.028*** (0.002)	0.040* (0.016)	—
Prior Incarceration (In Months)	0.062*** (0.000)	0.068*** (0.015)	—
Multiple Counts	0.095*** (0.003)	0.081* (0.031)	—
Trafficking Marijuana	−0.412*** (0.018)	−0.118 (0.071)	—
Trafficking Heroin	0.412** (0.018)	0.291 (0.101)	—
Trafficking Cocaine\Crack	0.540 (0.004)	0.313* (0.118)	—
Possess Marijuana	−0.386*** (0.005)	−0.214 (0.113)	—
Possess Heroin	−0.124* (0.013)	0.205 (0.104)	—
Possess Cocaine\Crack	−0.042*** (0.005)	0.150 (0.098)	—
Determinate Sentence	−0.300*** (0.003)	0.002 (0.112)	—
Community Correction Supervision	0.049** (0.003)	0.049 (0.037)	—
Year of Admission	−0.031*** (0.000)	−0.058*** (0.011)	—
Determinate Sentencing State	—	—	−0.602* (0.253)
Percent Below Poverty Level	—	—	0.154* (0.065)
Percent No High School Deg.	—	—	−0.127* (0.056)
Reliability Estimate	0.999	0.991	1.000
Level I R-Squared	0.084	0.260	—
Level II R-Squared	—	—	0.79
n (Level I)	144,263	144,263	144,263
n (Level II)	38	38	38
Chi-Square	188,432.19***	25,746.063***	174,966.822***

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

($P < 0.05$). The results of the fixed effects models indicate that African Americans have a 6% increase in sentence length compared to Caucasian offenders while controlling for other factors in fixed effect models. While all of the additive fixed slope models show that African Americans are predicted to have an increase in sentence length, there is a statistically significant random slope for this coefficient which renders the relationship insignificant. As such, the relationship varies across states due to some aggregate factor at level 2 (Snijders & Bosker, 2012). In model six and seven, cross-level interactions are explored in order to further elucidate this effect.

Given the importance of state variance in the assessment of the race and sentence length relationship, the researchers decided to test the effects of several state level indicators on this relationship in models 6 and 7. These indicate that the increase in sentence length is even greater when testing the effect of other level II factors on the treatment. Because of an insignificant ($P > 0.500$) chi-squared statistic for treatment when testing these effects, treatment was reset to fixed effects in these models. These models indicate that there is an increase in sentence length of about 20% when controlling for some of these effects in model 6. When testing for changes in the African-American population, the percent of marijuana arrests and use that are African American, the significance of the random slope disappears, and the statistical significance of the main effects of race reappears. This indicates that the random slope has been explained by the interaction with the level 2 covariates (Snijders & Bosker, 2012). Interestingly, Southern states have a decrease of

Table 4. Hierarchical linear models predicting log sentence length in months continued.

	Model IV Fixed	Model V Random	Model VI Fixed Treatment	Model VII Fixed Treatment
Treated (African American)	.056*** (.002)	.049 (.026)	.218*** (.029)	.057 (.060)
Southern State	—	—	−.053* (.004)	−.070** (.018)
Percent African American	—	—	.01** (.003)	.006* (.002)
Marijuana arrests % Black	—	—	−.007*** (.001)	−.003** (.001)
Marijuana use % Black	—	—	−.011* (.001)	.001*** (.001)
Determinate sentencing state	—	—	—	.033*** (.003)
Percent below poverty level	—	—	—	−.015* (.005)
Percent no high school deg.	—	—	—	.026*** (.003)
Age at admission	.002*** (.000)	.004*** (.001)	.004** (.001)	.003* (.011)
Male	.164*** (.002)	.165*** (.027)	.166*** (.028)	.165*** (.028)
Education	.019*** (.003)	.019** (.007)	.020* (.007)	.020* (.007)
Prior felony conviction	.028*** (.003)	.040* (.016)	.040 (.016)	.038 (.016)
Prior incarceration (in months)	.062*** (.000)	.077*** (.020)	.060*** (.014)	.064*** (.015)
Multiple counts	.095*** (.003)	.091** (.031)	.082* (.031)	.081* (.031)
Trafficking marijuana	−.160*** (.004)	−.192* (.079)	−.196 (..082)	−.212 (.079)
Trafficking heroin	.412* (.018)	.303** (.097)	.263 (.103)	.269 (.101)
Trafficking cocaine\crack	.531***(.004)	.337** (.113)	.269 (.122)	.272* (.119)
Possess marijuana	−.379** (.005)	−.215* (.095)	−.143 (.107)	−.159 (.105)
Possess heroin	−.123* (.013)	.138 (.088)	.193 (.101)	.167 (.100)
Possess cocaine\crack	−.044 (.005)	.068 (.080)	.169 (.099)	.121 (.090)
Determinate sentence	−.300*** (.003)	.009** (.104)	.029 (.107)	.004 (.107)
Community correction supervision	.048** (.003)	.049 (.036)	.048 (.036)	.051 (.036)
Year of admission	−.031*** (.001)	−.060*** (.011)	−.058*** (.011)	−.059*** (.011)
Determinate sentencing state	−.399 (.264)	−.392*** (.054)	−.413** (.138)	−.425*** (.044)
Percent below poverty level	.142* (.059)	.055* (.019)	.062 (.033)	.120 (.119)
Percent no high school deg.	−.135* (.052)	−.073*** (.012)	−.070 (.027)	−.061 (.061)
Southern state	—	−.456* (.198)	−.710 (.329)	−.563 (.366)
Population percent Black	—	—	.017 (.016)	.016 (.016)
Marijuana arrests % Black	—	—	.001 (.001)	.001 (.001)
Marijuana use % Black	—	—	.035 (.095)	−.036 (.101)
Reliability estimate	0.998	0.999	0.999	1.000
Level I R-squared	0.15	0.26	0.21	0.26
Level II R-squared	0.80	0.95	0.86	0.98
n (Level I)	144,263	144,263	144,263	144,263
n (Level II)	38	38	38	38
Chi-Square	181107.150***	4119.886***	3758.609***	6396.091***

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

about 5 to 7% percent in African-American sentence length compared to other states. The percent of marijuana arrests in a state that are African American leads to a small, yet significant decrease in the percent change in the treatment effect on sentence length. Furthermore, the percent of individuals that indicated that they had ever used marijuana that were African American resulted in about a 1% decrease in sentence length for African Americans.

Model 7 assesses the impact of poverty and education levels as well as determinate sentencing on the effect of race on sentence length. The introduction of these indicators renders the treatment effect of race an insignificant predictor of sentence length. While poverty level decreases the effect of race by about 2%, education increases the effect of race by about three percent. Finally, states with determinate sentencing had a three percent point increase in the effect of being African American on sentence length.

Several of the sociodemographic and criminal history control variables were significant predictors of the percent change in sentence length. The age of the offender was indicative of less than 1% change in sentence length per increase in age. Males had a 17% increase in sentence length. Each increase in the ordinal education scale was indicative of a 2% increase in sentence length. Furthermore, offenders who were under community corrections supervision prior to their incarceration had a 5% increase in sentence length; however, this finding was also limited to a few fixed effects models. Inmates with a prior felony conviction had 3 to 4% increase in sentence length compared to first offenders within the sample. Each percentage increase in months of prior incarceration was indicative of a 6 to 8% increase in the average sentence length.

Control indicators related to the offense the offender was sentenced to prison for are important and statistically significant predictors of the percent change in the length of sentence. Offenders charged with multiple counts had an 8 to 9% increase in sentence length compared to those with single counts. Determinate sentences resulted in a decreased change in sentence length of about 29% in fixed effect models. Each increase in the year of admission resulted in a 3 to 6% decrease in sentence length. Generally, persons incarcerated for trafficking marijuana had a decreased percentage change in sentence length when compared to those incarcerated for trafficking other drugs while offenders who were convicted of trafficking heroin or cocaine/crack had an increased percentage change in sentence length compared to others. Finally, statistical significance for possession offenses varied between models; however, those imprisoned for possession of marijuana, heroin, and cocaine generally had a decreased percentage change in sentence length compared to those incarcerated for possession of other drugs.

Several of the level 2 variables were significant predictors of the percentage change in sentence length. In model III determinate sentencing state, the percentage below the poverty level, and the percentage of the population with no high school degree were the only significant predictors; thus, others were removed from the model in order to increase statistical power and degrees of freedom. The results indicate that states that have determinate sentencing have a 60% decrease in the average sentence length. The percentage of the population living below the poverty level is indicative of a 20% increase in sentence length. The percent of the population without a high school degree yields a 13% decrease in sentence length. These percentages of change and significance vary slightly over models 4 through 6. Finally, Southern states have a 62% decrease in the average sentence length compared to other areas of the country.

Discussion and conclusion

The results of this study proffer a great deal of insight into the complexities of racial disadvantage within sentencing decisions across several dimensions. First, this study sheds light on incongruous findings within the literature while optimally

controlling for race using a quasi-experimental approach. The significant relationship between race and sentence length in the fixed models correspond with much of the previous literature (e.g., Arvanites, 2014; Crutchfield et al., 2010; Harris et al., 2009; Mitchell, 2005; Pratt, 1998; Schlesinger, 2011; Sorensen et al., 2003; Steffensmeier et al., 1998). This relationship persists despite the introduction of legal controls including prior convictions, the number of counts, and the type of offense, among others. The inability of legal predictors to explain the importance of race suggests that racial disparities exist beyond differential offending patterns, conflicting with the works of Barnes & Kingsnorth (1996), Crutchfield et al. (2010), Harris et al. (2009), Kleck (1981), Pratt (1998), and Sorensen et al. (2003). However, this relationship is only significant within the fixed models. Once race is allowed to vary, the importance of race disappears, suggesting that variation exists. The nonsignificant findings evident in the random effects model are compatible with the work of Jordan & Freiburger (2015). Relying on analyses that do not acknowledge varying means obscures some of the intricacies of this phenomenon. Therefore, variations in sampling and analytical strategies may, in fact, explain inconsistencies within the literature.

Second, causal explanations of this inequality are delineated by disentangling several interactions at the aggregate level. The results of this study are particularly compatible with the race-neutral interactions theories. The cross-level interaction of race and other variables in model seven supports this contention. Specifically, the main effects of race and sentencing are rendered nonsignificant when cross-level interactions are introduced for determinate sentencing and socioeconomic factors. Thus, the relationship between race and sentencing may be indirect in that it operates through socioeconomic inequality. This relationship between poverty level and sentencing outcome is consistent with research finding an inverse relationship between measures of welfare and inequality and incarceration (see Beckett & Western, 2004).

Though the results of this study are supportive of the interaction argument, some of the contextual level interactions are not necessarily race-neutral. This finding could be construed as supportive of a direct effects argument, an argument alluded to in numerous previous studies (see Barnes & Kingsnorth, 1996; Blumstein, 1982; Crutchfield et al., 2010; Harris et al., 2009; Kleck, 1981; Pratt, 1998; Sorensen et al., 2003). For instance, the relationship between the percentage of African Americans per state and sentence lengths is consistent with Hawkins and Hardy's (2015) finding that inequality is greater in states with smaller African-American populations. Furthermore, both legal and extralegal factors fail to entirely explain the race gap in sentence length at level 1, a finding incongruent with the race-neutral interactions perspective. Racial differences continue to persist when controlling for important extralegal factors such as age, gender, and education, challenging the assertion that racial disparities are a product of race-neutral variables. However, these race-based aggregate interactions moderate the relationship between race and sentencing rather than mediating it. The introduction of

these factors greatly increased the effects of race on sentencing (from 5% to 22%). Moreover, others have argued that differential involvement and arrest rates have been found to merely mitigate regional variation in racial sentencing disparities but not alleviate its effects (Sorenson et al., 2003). Similarly, increases in the proportion of marijuana users and arrestees that are African American within a state is related to a reduction in the predicted increase in sentence length for African Americans.

Third, the hypothesized geographically specific prejudicial treatment of African Americans within the South is contested. Contrary to expectation, the Southern states did not appear to demonstrate particularly punitive sentencing outcomes toward African Americans compared to other regions (see Oshinsky, 1997). Though African Americans received longer sentences than Caucasians overall, the race gap was smaller in the South than in other regions. Unlike Kleck's (1981) study, the results of this study suggest that racial discrimination within Southern courts may not be as prominent as assumed, especially compared to other regions, despite a sociohistorical background that might suggest otherwise. This may be a result of changes in population characteristics and sentiments in the South over time. Furthermore, in recent decades there have been substantial increases in minority decision makers such as judges and prosecutors (Ward, Farrell, & Rousseau, 2009). While this rise in nonwhite courtroom decision makers is evident across the United States, this movement may be of particular importance within the South as a product of larger African-American populations. Nonetheless, the smaller sentencing disparities in the South are consistent with other studies that have found that Midwestern states appear to be more disproportionately punitive towards African Americans regarding imprisonment than the South (Blumstein, 1993; Christianson, 1981; Tonry, 1991).

Fourth, this paper challenges the utility of determinate sentencing strategies as a mechanism for reducing inequality. It appears as though determinate sentencing, rather than removing racial bias in sentencing, increases sentence length for African Americans. Some have argued that these laws exacerbate inequality rather than alleviating it (Alexander, 2010). Similarly, determinate sentencing states have been found to be related to increased sentence length for African-American offenders and increased racial disproportionate prison admissions (Schlesinger, 2011). This may be a result of changes in the goals of determinate sentencing over time (Mauer, 2001; Travis, Western, & Redburn, 2014). Specifically, the war on drugs has shifted the focus of determinate sentencing away from limiting discrimination to increasing the certainty and severity of punishment through legislation such as mandatory minimum, three strikes, and truth in sentencing legislation (Mauer, 2001; Travis et al., 2014). Contrarily, the findings indicate that overall determinate sentencing leads to a reduction in sentence length. Thus, determinate sentencing may lead to decreases in sentence length for some while leading to increases for others.

Determinate sentencing legislation has greatly reduced the discretion of the judge and jury while greatly expanding the power and discretion of the prosecutor (Simon, 2007). Additionally, Zatz (1987) argued that racial discrimination may be a product of prosecutorial gatekeeping given the wide discretion available to these courtroom decision makers. Thus, the interactional relationship between determinate sentencing and sentencing outcomes for the treatment group may be a result of prosecutorial decision making rather than that of the judge or jury. The discretion of the prosecutor is important when considering the complexity and variations in the criminal code both across and within states. Furthermore, many states distinguish between simple possession of a drug and possession with intent to distribute a substance. The prosecutor's ability to decide the criminal charge has great impact on the sentence because the sentence varies greatly with the charge (e.g., some may carry a mandatory minimum, while others do not). The interaction of socioeconomic factors with race may be important here as well, in that those with the economic resources to hire an attorney will have increased ability to influence the decisions of the other important actors. The importance of discretion is illustrated in the increased sentence length prediction for offenders charged under determinate sentencing statutes while controlling for determinate sentencing states in model 5 as well as its interaction with race in model 7.

While these intriguing findings have the potential to contribute toward the current literature on sentencing and inequality, some limitations are discernable within this study. The absence of excluded cases could contribute to sampling bias as a result of heterogeneous attrition. However, the utility of the propensity score matching outweighed the detrimental cost of losing cases. Moreover, due to the propensity score matching method utilized herein, there was not sufficient data within counties to assess the variance explained at this level. The combination of crack and cocaine offenses into one category within these data may disguise some of the racial variations in sentencing patterns (see, e.g., Lowney, 1994; Chappell & Maggard, 2007). Finally, these data were limited to the 38 states that report to the NCRP; thus we cannot generalize this to the entire county as the sample selection was nonrandom.

Furthermore, though the results of this study offer moderate support for the interactional and differential response arguments, caution should be exercised when deciding the legitimacy of these theoretical explanations. Ideally, additional controls would be implemented in order to further establish causal relationships and alleviate the potential for spurious effects. For instance, the legitimacy of the differential offending explanation cannot be entirely determined due to the omission of several key legal variables, including the presence of a weapon, whether the case went to court, and the amount of drugs involved to name a few. As with the differential offending argument, race-neutral explanations cannot be accepted without question without further examining level 1 variations in SES, the number of dependents, citizenship status, as well as other extralegal factors. Finally, direct discrimination should not be dismissed without first collecting data from decision

makers including prosecutors and judges. Unfortunately, none of the variables identified above were collected by the National Corrections Reporting Program. Future research should examine additional courtroom decisions in addition to sentence length. Decisions to incarcerate as well as the application of alternative sanctions may demonstrate disparities unique from those related to sentence length. In order to examine potential direct effects on racial disparities, a mixed-method approach using interviews from judges and prosecutors would be of particular importance.

Despite these limitations, this project contributes to the existing literature by more fully controlling for legal and extralegal differences between races using propensity score matching. Additionally, the numerous multilevel models are able to illustrate support, or lack thereof, for many of the theories of racial disproportionality in sentencing, depending on the model. Thus, the mixed results of some of the prior research are explained by the differences in the models and their results. For example, a simple fixed effects model would have been supportive of a direct impact of racial discrimination. However, further analysis demonstrates that many aggregate social, economic, and demographic interactions play a meaningful role in the race and sentencing relationship. Moreover, this endeavor raises some interesting quandaries about the effectiveness of determinate sentencing and challenges previously held notions of the prejudicial South. In conclusion, it appears racial and/or social inequality is, in fact, present within drug sentencing decisions; however, this relationship is not entirely black and white.

Notes on contributors

Richard J. Stringer is a Ph.D. candidate in Criminology and Criminal Justice at Old Dominion University in Norfolk, Virginia. His research and teaching interests include substance use, drug and alcohol policy, alcohol impaired driving, the media and public opinion, and social problems. His research has appeared in *Criminal Justice Policy Review, Journal of Drug Issues*, and the *Journal of Ethnicity in Criminal Justice.*

Dr. Melanie M. Holland is an assistant professor in the department of Criminal Justice at Pennsylvania State University in Altoona, Pennsylvania. She received her Ph.D. in Criminology and Criminal Justice at Old Dominion University in Norfolk, Virginia. Her research and teaching interests include sentencing, social inequality, and race and ethnicity. Her work has appeared in *Criminal Justice Review* and the *Journal of Ethnicity in Criminal Justice*.

References

Alexander, M. (2010). *The New Jim Crow: Mass incarceration in the age of colorblindness*. New York, NY: The New Press.

Allison, P. D. (2001). *Missing data* (Vol. 136). Thousand Oaks, CA: Sage Publications.

Arvanites, T. (2014). Segregation and African-American imprisonment rates for drug offenses. *The Social Science Journal*, *51* (3), 431–437.

Arvanites, T. M., & Asher, M. A. (1995). The direct and indirect effects of socio-economic variables on state imprisonment rates. *Criminal Justice Policy Review*, *7* (1), 27–53.

Barnes, C. W., & Kingsnorth, R. (1996). Race, drug, and criminal sentencing: Hidden effects of the criminal law. *Journal of Criminal Justice*, *24* (1), 39–55.

Beckett, K., & Western, B. (2001). Governing social marginality welfare, incarceration, and the transformation of state policy. *Punishment & Society*, *3* (1), 43–59.

Blumstein, A. (1982). On the racial disproportionality of United States' prison populations. *Journal of Criminal Law & Criminology*, *73* (3), 1259–1281.

Blumstein, A. (1993). Racial disproportionality of U.S. prison populations revisited. *U. Colo. l. Rev.*, *64*, 743.

Bourassa, S. C., & Viviana, A. (2009). Decomposition of racial differences in sentencing: Application of an econometric technique to cocaine possession cases. *Journal of Ethnicity in Criminal Justice*, *7* (3), 186–204.

Chappell, A. T., & Maggard, S. R. (2007). Applying Black's theory of law to crack and cocaine dispositions. *International Journal of Offender Therapy and Comparative Criminology*, *51* (3), 264–278.

Christianson, S. (1981). Our black prisons. *Crime & Delinquency*, *27* (3), 364–375.

Crutchfield, R. D., Fernandes, A., & Martinez, J. (2010). Racial and ethnic disparity and criminal justice: How much is too much. *Journal of Criminal Law & Criminology*, *100* (3), 903–932.

Everett, R. S., & Wojtkiewicz, R. A. (2002). Difference, disparity, and race/ethnic bias in federal sentencing. *Journal of Quantitative Criminology*, *18* (2), 189–211.

Federal Bureau of Investigation. (2013a). *Number and Rates of Arrests by Region, 2013*. Retrieved from http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013/tables/table-30/table_30_number_and_rate_of_arrests_by_region_2013.xls

Federal Bureau of Investigation. (2013b). *Arrests for Drug Abuse Violations Percent Distribution by Region, 2013*. Retrieved from http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013/persons-arrested/persons-arrested

Federal Bureau of Investigation. (2013c). *Arrests by Race, 2013*. Retrieved from http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013/tables/table-43

Gerber, R. J. (2004). *Legalizing marijuana: Drug policy reform and prohibition politics*. Westport, CT: Greenwood Publishing Group.

Golub, A., Johnson, B. D., & Dunlap, E. (2007). The race/ethnicity disparity in misdemeanor marijuana arrests in New York City. *Criminology & Public Policy*, *6* (1), 131.

Guo, S., & Fraser, M. (2014). *Propensity score analysis: Statistical methods and applications* (Vol. 11). Thousand Oaks, CA: Sage Publications.

Harris, C. T., Steffensmeier, D., Ulmer, J. T., & Painter-Davis, N. (2009). Are Blacks and Hispanics disproportionately incarcerated relative to their arrests? Racial and ethnic disproportionality between arrest and incarceration. *Race and social problems*, *1* (4), 187–199.

Hawkins, D. F., & Hardy, K. A. (1989). Black-white imprisonment rates: A state-by-state analysis. *Social Justice*, *16* (4), 75–94.

Inter-University Consortium for Political and Social Research (ICPSR). (2014). National Corrections Reporting Program, 2000–2012 Codebook. Ann Arbor, MI.

Jordan, K. L., & Freiburger, T. L. (2015). The effect of race/ethnicity on sentencing: Examining sentence type, jail length, and prison length. *Journal of Ethnicity in Criminal Justice*, *13* (3), 179–196.

Katznelson, I. (2005). *When affirmative action was white: An untold history of racial inequality in twentieth-century America*. New York, NY: W. W. Norton & Company.

Kim, J.-S., & Steiner, P. M. (2015). Multilevel propensity score methods for estimating causal effects: A latent class modeling strategy. *Quantitative Psychology Research* (pp. 293–306). Madison, WI: Springer.

Kleck, G. (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.

Lowney, K. (1994). Smoked not snorted: Is racism inherent in our crack cocaine laws? *Urban Law Annual; Journal of Urban and Contemporary Law, 45*, 121.

Lyons, T., Lurigio, A. J., Roque, L., & Rodriguez, P. (2013). Racial disproportionality in the criminal justice system for drug offenses a state legislative response to the problem. *Race and Justice*, *3* (1), 83–101.

Mauer, M. (2001). The causes and consequences of prison growth in the United States. *Punishment & Society*, *3* (1), 9–20.

Massey, D. (2007). *Categorically unequal: The American stratification system*. New York, NY: Russel Sage Foundation.

Michalowski, R. J., & Pearson, M. A. (1990). Punishment and social structure at the state level: A cross-sectional comparison of 1970 and 1980. *Journal of Research in Crime and Delinquency*, *27* (1), 52–78.

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

Musto, D. (1999). *The American disease: Origins of narcotic control* (3rd ed.). New York, NY: Oxford University Press.

Oshinsky, D. (1997). *Worse than slavery*. New York, NY: Simon and Schuster.

Pasko, L. (2002). Villain or victim: Regional variation and ethnic disparity in federal drug offense sentencing. *Criminal Justice Policy Review*, *13* (4), 307–328.

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

Pritchard, A., & Wiatrowski, M. (2008). Race and capital punishment: State level analysis of the effects of race on states' capital punishment policies. *Journal of Ethnicity in Criminal Justice*, *6* (2), 103–121.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks, CA: Sage Publications.

Schlesinger, T. (2011). The failure of race neutral policies: How mandatory terms and sentencing enhancements contribute to mass racialized incarceration. *Crime & Delinquency*, *57* (1), 56–81.

Schoenfeld, H. (2012). War on drugs, the politics of crime, and mass incarceration in the United States. *Journal of Gender Race & Justice*, *15* (2), 315–352.

Simon, J. (2007). Governing through crime: How the war on crime transformed American democracy and created a culture of fear. New York, NY: Oxford University Press.

Snijders, T. A., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods & Research*, *22* (3), 342–363.

Snijders, T. A., & Bosker, R. J. (2004). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage Publications.

Sorensen, J., Hope, R., & Stemen, D. (2003). Racial disproportionality in state prison admissions: Can regional variation be explained by differential arrest rates? *Journal of Criminal Justice*, *31* (1), 73–84.

Steiner, P., Kim, J., & Thoemmes, F. (2013). *Matching strategies for observational multilevel data*. Paper presented at the Joint Statistical Meeting of the American Statistical Association proceedings.

Steffensmeier, D., Ulmer, J., & Kramer, J. (1998). The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black, and male. *Criminology*, *36* (4), 763–798.

Stolzenberg, L., D'Alessio, S. J., & Eitle, D. (2013). Race and cumulative discrimination in the prosecution of criminal defendants. *Race and Justice*, *3* (4), 275–299.

Taylor, M. C. (1998). How white attitudes vary with the racial composition of local populations: Numbers count. *American Sociological Review*, *63* (4), 512–535.

Thoemmes, F. J., & West, S. G. (2011). The use of propensity scores for nonrandomized designs with clustered data. *Multivariate Behavioral Research*, *46* (3), 514–543.

Tonry, M. (1991). Black, White incarceration rates. *Overcrowded Times*, *2* (3), 5–7.

Travis, J., Western, B., & Redburn, S. (2014). The growth of incarceration in the United States. Washington, D.C.: The National Academies Press.

Ulmer, J., & Light, M. T. (2011). Beyond disparity: Changes in federal sentencing after Booker and Gall? *Sentencing and Social Science*, *23* (5), 333–341.

United States Census Bureau. (2010). United States Census 2010. Retrieved from http://www.census.gov/2010census/data/

United States Census Bureau. (2010a). Geographic terms and concepts-Census divisions and Census regions. Retrieved from https://www.census.gov/geo/reference/gtc/gtc_census_divreg.html

United States Census Bureau. (2013). *State and County QuickFacts*. Retrieved from http://quickfacts.census.gov/qfd/states/00000.html

United States Department of Health and Human Services. (2011). Results from the 2010 National Survey on Drug Use and Health: Summary of national findings. Substance Abuse and Mental Health Services Administration: Rockville, MD, USA.

United States Department of Health and Human Services. (2013). Substance Abuse and Mental Health Services Administration. Center for Behavioral Health Statistics and Quality. National Survey on Drug Use and Health: 10-Year Substate R-DAS (2002 to 2011). ICPSR34947-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2013-12-17. doi:10.3886/ICPSR34947.v1

United States Department of Justice. Federal Bureau of Investigation. (2011). Uniform Crime Reporting Program Data: National Incident-Based Reporting System. ICPSR34585-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2013-09-05. doi:10.3886/ICPSR34585.v1

United States Department of Justice. (2014). Office of Justice Programs. Bureau of Justice, S. National Corrections Reporting Program, 2000–2012. Retrieved from http://doi.org/10.3886/ICPSR34984.v1

Ward, G., Farrell, A., & Rousseau, D. (2009). Does racial balance in workforce representation yield equal justice? Race relations of sentencing in federal court organizations. *Law and Society Review*, *43* (4), 757–806.

Weitzer, R. (1996). Racial discrimination in the criminal justice system: Findings and problems in the literature. *Journal of Criminal Justice*, *24* (4), 309–322.

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

Zatz, M. S. (1987). The changing forms of racial/ethnic biases in sentencing. *Journal of Research in Crime and Delinquency*, *24* (1), 69–92.