The University Of Sheffield.

This is a repository copy of *Conditional race disparities in criminal sentencing: a test of the liberation hypothesis from a non-guidelines state*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/93378/

Version: Accepted Version

Article:
Hester, R. and Hartman, T.K. (2017) Conditional race disparities in criminal sentencing: a test of the liberation hypothesis from a non-guidelines state. Journal of Quantitative Criminology, 33 (1). pp. 77-100. ISSN 0748-4518

https://doi.org/10.1007/s10940-016-9283-z

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

CONDITIONAL RACE DISPARITIES IN CRIMINAL SENTENCING: A TEST OF THE LIBERATION HYPOTHESIS FROM A NON-GUIDELINES STATE

Suggested Running Head: Conditional Race Disparities

Rhys Hester[1]

Todd K. Hartman[2]

[1]Robina Institute of Criminal Law and Criminal Justice, University of Minnesota Law School, 229 19th Avenue South, Minneapolis, MN 55455, USA. Email: rahester@umn.edu; phone: (612) 625-6146.

[2]Sheffield Methods Institute, The University of Sheffield, Sheffield, UK.

ABSTRACT

Objectives: To test the liberation hypothesis in a judicial context unconstrained by sentencing guidelines.

Methods: We examined cross-sectional sentencing data (n = 17,671) using a hurdle count model, which combines a binary (logistic regression) model to predict zero counts and a zero-truncated negative binomial model to predict positive counts. We also conducted a series of Monte Carlo simulations to demonstrate that the hurdle count model provides unbiased estimates of our sentencing data and outperforms alternative approaches.

Results: For the liberation hypothesis, results of the interaction terms for race x offense severity and race x criminal history varied by decision type. For the in/out decision, criminal history moderated the effects of race: among offenders with less extensive criminal histories blacks were more likely to be incarcerated; among offenders with higher criminal histories this race effect disappeared. The race x offense severity interaction was not significant for the in/out decision. For the sentence length decision, offense severity moderated the effects of race: among offenders convicted of less serious crimes blacks received longer sentences than whites; among offenders convicted of crimes falling in the most serious offense categories the race effect became non-significant for Felony D offenses and transitioned to a relative reduction for blacks for the most serious Felony A, B, and C categories. The race x criminal history interaction was not significant for the length decision.

Conclusions: There is some support for the liberation hypothesis in this test from a non-guidelines jurisdiction. The findings suggest, however, that the decision to incarcerate and the sentence length decision may employ different processes in which the interactions between race and seriousness measures vary.

KEYWORDS

Sentencing; Racial Disparities; Criminal Sentencing; Liberation Hypothesis; Hurdle Model

1. INTRODUCTION

Recent assessments of the criminal sentencing literature have noted two important goals: (1) identifying the extent to which racial disparities are present in the judicial process, and (2) examining the effects of various policy initiatives on sentencing outcomes (Baumer, 2013; Ulmer, 2012). While much progress has been made toward these objectives, scholars such as Baumer (2013) and Ulmer (2012) continue to highlight the need for studies that examine not just whether race matters in sentencing, but also how and when race factors into judicial decision-making (see also Spohn, 2000). These calls are reinforced by a growing literature that finds racial disparities in certain parts of the criminal justice process but not others (e.g., Blumstein, 1982, Kutateladze et al., 2014; Rehavi and Starr, 2014). Moreover, there has been a particular emphasis on the need to examine sentencing practices in a broader variety of contexts—in places other than guidelines jurisdictions like Minnesota, Pennsylvania, Washington, and the federal system, which have dominated the literature (Engen, 2009; Reitz, 2009; Ulmer, 2012).

The current study investigates the extent to which the severity of the offense and the prior record of the offender condition the likelihood that black offenders receive more punitive treatment in a jurisdiction unconstrained by sentencing guidelines. Applied to criminal sentencing, the "liberation hypothesis" posits that judicial decision-makers will feel constrained to sentence offenders in an equally harsh manner in the most serious cases (Kalven and Zeisel, 1966; Spohn and Cederblom, 1991). In such instances, extralegal characteristics like race of the offender will not be considered given the overshadowing importance of factors like offense severity and prior record. However, for defendants with less severe cases and less extensive criminal histories, greater ambiguity surrounds the sentencing decision; thus, judges will feel "liberated" to individualize the sentence on a variety of factors. This ambiguity increases the likelihood that sentencing decisions might be influenced by the race of the offender (Spohn, 2000; Spohn and Cederblom, 1991; Spohn and DeLone, 2000).

We offer several contributions to the study of criminal sentencing by examining potential conditioning effects of offense seriousness and criminal history on race. We provide a robust test of the liberation hypothesis using data from 17,671 criminal offenders in the state of South Carolina. This state is particularly interesting for these purposes because there are no sentencing guidelines, which means that decision-makers have greater discretion when sentencing offenders. Methodologically, we employ a class of event count models, which better handle positively skewed distributions such as those found in sentencing data thereby allowing the researcher to fit a model to the data rather than manipulating the data to fit a model. In sum, testing the interaction of certain offender characteristics with offense severity and prior record may provide insight into when disparities manifest in sentencing decisions.

2. PRIOR LITERATURE

The primary “meta-goal” of sentencing research has been to explore racial disparities in sentencing outcomes (Baumer, 2013; see also Spohn, 2000; Zatz, 2000). In explaining racial disparities in sentencing, social science researchers have generally relied on theoretical explanations rooted in symbolic interactionism, a sociological theory which holds that an actor’s words and actions toward another entity are based on meanings the actor ascribes to the other person, event, situation, or thing (Blumer, 1969; Ulmer, 1997; Wooldredge, 2007). In the court context, symbolic interactionism suggests that judicial decision-making is a function of the meaning ascribed to an offender’s characteristics, actions, and past behaviors—for example, the meaning a judge gives to a “young person,” a “black man,” a “violent offender,” or a “repeat offender.” More specifically, courtroom actors develop patterned responses to certain cues such as the seriousness of the offense, whether it involved violence, and the defendant's criminal record, as well as extralegal characteristics like race, gender, and socioeconomic status (Albonetti, 1991; Steffensmeier, Ulmer, and Kramer, 1998).

For Steffensmeier and colleagues (1998) these symbolic interactions coalesce around three focal concerns of sentencing. Courtroom decision-makers: (1) emphasize blameworthiness by imposing severe penalties on offenders who commit serious offenses, who have extensive criminal histories, or who cause more harm; (2) seek to protect the community from dangerous offenders by attempting to anticipate future behavior (which calls on the decision-maker's attributions about characteristics of the case and offender); and (3) attempt to navigate practical constraints such as the availability of jail space and the desire to maintain relationships in the courtroom workgroup. Drawing on the work of Albonetti (1991) and others, Steffensmeier and colleagues note that the decisions informed by these three focal concerns are based on limited information and involve uncertain predictions—for example, about who is likely to reoffend. Accordingly, court actors develop "perceptual shorthands" which appear to include salient stereotypes for young minority males.

Several systematic reviews of the sentencing literature have concluded that race effects are often present in the decision to incarcerate, though much less likely to occur for the length of the sentencing decision (Chiricos and Crawford, 1995; Mitchell, 2005; Spohn, 2000). Chiricos and Crawford (1995) found more pronounced disparities in Southern jurisdictions, though region was not a statistically significant predictor of the relationship between race and sentencing outcomes for Mitchell (2005). Based on these reviews, which cover hundreds of studies, Baumer (2013: 242) concluded that there are "small but statistically significant direct race differences in the probability of imprisonment to the disadvantage of blacks," but little evidence of "direct race differences in prison sentence lengths between these two groups."

2.1 Interaction Effects and the Liberation Hypothesis

Increasingly, scholars have shifted attention away from direct effects to investigate conditional effects of race and other characteristics on sentencing decisions. Spohn (2000: 432), for instance, intimated "it is overly simplistic to assume that minorities will receive harsher sentences than whites regardless of the crime, the seriousness of the offense, or the culpability of the defendant." Both Spohn (2000) and Zatz (2000) suggest a focus on ways that race might conditionally affect sentencing by modeling interactions (see also Spohn and DeLone, 2000). One line of contextualized disparity inquiry—and our focus for this paper—concerns the liberation hypothesis. The liberation hypothesis was originally articulated by Kalven and Zeisel (1966) in their landmark study of jury decision-making in which they used various case characteristics such as the evidence presented and demographic profile of the defendant to predict a jury's likelihood of coming to a "correct" decision. The key factor that determined whether demographic attributes would factor into the decision was the strength of the evidence in the case, or as Kalven and Zeisel (1966: 165) fashioned it, whether the evidence was weak enough to make the case a "close" one: "The closeness of the evidence makes it possible for the jury to respond to sentiment by liberating it from a discipline of the evidence." With a dearth of evidence a not-guilty verdict was apparent, as was a guilty verdict with an abundance of strong evidence. But for cases in between—the close evidence cases—extralegal factors such as juror views about the defendant factored into the decision-making process.

The liberation hypothesis has since been adapted and extended to examine decision-making in capital cases (Baldus, Woodworth, and Pulaski, 1990), police use of force (Barkan and

Cohn, 1998), prosecutorial decision-making (Ball, 2006), juvenile court decision making (Guevara et al., 2011), application of three strikes enhancements (Chen, 2008), and parole violation decisions (Grattet and Lin, 2014; Lin, Grattet, and Petersilia, 2012). A number of studies have also applied the hypothesis to the investigation of racial disparities in traditional sentencing outcomes of the disposition decision (in or out of prison) and the duration decision (length of incarceration)[3] (e.g., Lieber and Blowers, 2003; Spohn and Cederblom 1991; Spohn and DeLone, 2000; Warren, Chiricos, and Bales, 2012).

The liberation hypothesis holds theoretical appeal and has received renewed interest in recent years. However, the results from these studies have been extremely mixed (see, e.g., Spohn and Cederblom, 1991, finding support for the hypothesis for the disposition decision but not for the duration decision; Spohn and DeLone, 2000, reporting results consistent with the liberation hypothesis for some city- and minority-group combinations but not others; Warren, Chiricos, and Bales, 2012, finding support for the hypothesis among some offender/case type/outcome combinations but not others). Perhaps explanations for these null and even contrary results can be found in the newest wave of contextual disparity research (see, e.g., Kutateladez et al., 2014; Rehavi and Starr, 2014; Starr, 2015). First, not only might stage of the process affect the presence or absence of disparities, but factors like geographic context likely matter (see, e.g., Eisenstein et al., 1988; Kramer and Ulmer, 2009; Ulmer, 1997), and

[3] Several researchers have also suggested integrating focal concerns with the liberation hypothesis. Lieber and Blowers (2003) took a first step in their study of misdemeanor sentencing. They note that while the liberation hypothesis gives a reason why sentiments are more influential at lower levels of severity, the hypothesis does not offer an explanation why those sentiments might manifest as racial bias. Lieber and Blowers (2003) turn to focal concerns to bridge this gap. Ball (2006) also paired focal concerns with the liberation hypothesis in his study of prosecutorial plea bargaining, and Guevara et al., (2011) suggest further integration of focal concerns and attributional theory as a path for future research.

accordingly, we should anticipate null findings in some jurisdictions and the presence of racial disparities in others. Second, under certain case constellations, theoretical mechanisms operating outside the liberation paradigm (e.g., the stereotype of the young minority male for violent crimes) might counteract the more general tendency for disparities to be least when severity is greatest. As such, we should not expect to find racial disparities everywhere.

With the current study we offer an analytical contribution to the sentencing field by testing the liberation hypothesis in a well-suited jurisdictional context. First, we examine the entire range of offenses from serious misdemeanors to the most serious felonies—all offenses sentenced in the general jurisdiction courts for our population. Second, some prior studies have estimated separate regression models for different demographic characteristics which do not actually test whether variables of interest (such as offense seriousness and prior record) significantly differ by race. Finally, we test the liberation hypothesis in a jurisdiction where judges are not confined by sentencing guidelines.

2.2 Research Expectations

Following Bushway and Piehl (2001: 735), we consider "judicial discretion" to be "discretion in criminal sentencing regardless of whether a judge or prosecutor is responsible for setting a given sentence and regardless of whether the sentence results from a trial or a guilty plea." The judicial discretion enjoyed by courtroom workgroup actors in this study should have created an optimal context for race-based differences to manifest along different levels of offense severity and prior record. Based upon the liberation hypothesis, we expect judicial decision-makers to be more punitive towards blacks relative to whites when there is greater ambiguity surrounding the sentencing decision. For the more serious offenders, decisions will be dominated by the magnitude of the offense or record of the offender. But where offense severity and prior record are lower, the question of appropriate punishment becomes less clear, and opportunities for the influence of extralegal characteristics like race will likely increase, manifesting in greater disparity.

3. DATA AND METHOD

The data for this study consist of criminal cases sentenced in South Carolina Circuit Courts (the courts of general jurisdiction) for the fiscal year 2001. The South Carolina Sentencing Commission (now disbanded) compiled the data to facilitate the creation of advisory sentencing guidelines which were proposed to the S.C. General Assembly but never adopted. The Circuit Courts have jurisdiction over felony and serious misdemeanor offenses.[4] We chose to analyze all felony cases, as well as cases which South Carolina labeled misdemeanors, but which were serious offenses that might be labeled felonies in other jurisdictions. For instance, several South Carolina misdemeanors carried maximum penalties of several years in prison with some carrying up to 10 years in prison. During FY2001, aggravated assault, obviously a serious crime and one of the most common offenses in the data, was labeled by South Carolina law as an

[4] During FY2001, lower level magistrate courts had jurisdiction to sentence offenses subject to a maximum 30 days incarceration, a $500 fine, or both; or up to one year in prison, a $5,500 fine, or both upon transfer from the Circuit Court (S.C. Code §§ 22-3-550, -545). Criminal jurisdiction for all other cases rested with the Circuit Courts (S.C. Code § 14-25-65). The Commission data did not contain a record of all misdemeanor offenders sentenced in the lower courts. Because neither the complete population nor a representative sample of misdemeanor offenders was available, it was not possible to examine the universe of misdemeanor and felony sentencing outcomes. Accordingly, we included all felonies and serious misdemeanors carrying the potential for more than one-year incarceration, which is the traditional definition of a felony offense (McAninch, et al., 2007). This allowed us to include offenses which were deemed serious enough by the S.C. legislature to merit the potential for more than a year in prison, while also removing the unrepresentative portion of misdemeanor offenses that happened to have been sentenced in Circuit Court rather than a lower-level court. Classifying offenses this way also makes these analyses more comparable to the existing research on felony sentencing conducted in other states, rather than constituting a study marked by anomalous state law designations.

unclassified misdemeanor with a maximum penalty of 10 years in prison. Excluding such serious offenses solely due to the designation as misdemeanor rather than felony seemed inappropriate. Accordingly, we included all misdemeanors that met the traditional definition of a felony—that is, subject to a custody sentence of more than one year in prison.[5]

The original Commission data did not include whether offenders pled guilty or were sentenced after a trial. Because prior research has found mode of disposition to be a significant predictor of both the incarceration and sentence length decision (Kramer and Ulmer, 2009; Spohn, 2009), we supplemented the Commission data with the mode of disposition through a request to S.C. Court Administration.[6] The data also included 429 individuals who were sentenced twice in FY2001 and comprised two separate cases. In these instances, only the most serious offense entry was kept. The Commission data did not distinguish Hispanic offenders and apparently accounted for non-white and non-black offenders inconsistently. As such, we dropped 221 individuals whose race was entered as "other." Four cases were deleted for missing data on offense seriousness, and 2 were dropped for missing offender race. These delimitations resulted in a dataset of 17,671 offender cases, including 6,611 offenders who were incarcerated (i.e., sentenced to a prison term greater than 0). These cases represented all offenders convicted of a

[5] While misdemeanors with potential prison sentences of more than one year were included as the lowest offense severity level, we recoded the unclassified common law offenses which were subject to 10 year maximums as Class E felonies because Class E felonies were capped at a maximum of 10 years (S.C. Code Ann. Sections 17-25-20, 17-25-30; McAninch, Fairey, and Coggiola 2007).

[6] Starting with the supplemental Court Administration list of all criminal cases that went to trial in FY2001, we successfully matched 85% of these (260 of 306 total trials) with the Commission data. Some of the failed matches were sealed cases listed in the supplemental Court Administration data that might have been excluded from the Commission's dataset, while other cases failed to match for unknown reasons.

felony (or serious misdemeanor carrying a maximum of a year or more in prison) who were sentenced by active Circuit Judges in the general jurisdiction courts for FY2001.

3.1 Measures

A brief description of the measures, as well as summary statistics, is provided in Table 1. Our primary dependent variable, prison term, is an offender's expected minimum sentence, which helps account for the non-uniform nature of parole eligibility for South Carolina offenses (Freiberger and Hilinski, 2013; see also Chiricos and Bales, 1991; Spohn and Cederblom, 1991).[7] [8] Because the data include extreme outliers such as life and death sentences, we top coded 0.44% of the cases at a maximum of 720 months, or an expected minimum sentence of 60 years in prison.[9]

---

[7] The expected minimum sentence was chosen over other alternatives because offenders might have been eligible for parole after serving 25 percent, 33 percent, or 85 percent of their sentences, or may never have been eligible, depending upon the classification of the offense. Using the expected minimum rather than the imposed maximum accounted for these differences in parole eligibility (Chiricos and Bales, 1991; Gertz and Price, 1985; Spohn and Cederblom, 1991). The expected minimum was calculated by adjusting the imposed maximum sentence by a parole eligibility multiplier as determined by the controlling offense (e.g., 0.25, 0.33, 0.85, 1.0) and rounded up to the nearest month (less than 200 of the 17,671 original sentences are non-integers). For example, if an offender was sentenced to 10 years and fell under the 25 percent parole eligibility designation, the expected minimum would be 2.5 years (10 x .25), or 30 months.

[8] We were not able to discern between prison and jail sentences. Defendants sentenced to more than three months custodial time are processed into the state correctional system; defendants given less than three months serve their time in local jails or detention centers. Thus, unlike in some states, Circuit Court judges do not make an independent decision whether to send incarcerated offenders to a local jail or central prison—that decision is a product of the length of the sentence imposed.

[9] Choosing a cut-point for the top coding is somewhat arbitrary, and some scholars have used other operationalizations (e.g., Johnson, Ulmer, and Kramer, 2008, top coded at 470 months based on the federal sentencing commission's convention of using 470 months as representative of a life sentence). Our findings were robust to other coding decisions. For example, we ran supplemental models that altered the value of right-censoring, omitted life and death sentences altogether, and used the unmodified (raw) sentence, and our substantive findings were not meaningfully different across models.

We also included case and offender measures generally found in sentencing research. Offense seriousness is an ordinal measure of the severity of a crime based upon the S.C. Crime Classification Scheme. Offenses were coded '1' for misdemeanors carrying a possible sentence of over a year in prison, '2' for Class F Felonies, '3' for Class E Felonies, '4' for Class D Felonies, and '5' for Class A, B, or C Felonies (or Unclassified Felonies). The S.C. Sentencing Commission also created a measure of an offender's criminal history, for which 4 points were assigned for each prior violent or drug trafficking conviction with a sentence of a year or greater, 2 points for prior sentences of less than a year, and 1 point (up to a maximum of five) for prior non-incarceration convictions. Offenders with a score of '0' were deemed to have had no prior criminal history, while those with '1 to 3' had minimal, '4-12' had moderate, '13-20' had considerable, and '21 and over' had extensive criminal histories.

Commitment score is an ordinal measure based upon the number and severity of the offenses for which one was currently found guilty. All offenders received 1 point for their main offense. Beyond that, offenders were given 1 point for each additional count or offense unless any of those additional offenses were for an A, B, C, or Exempt Felony offense; in these instances, 4 points were added to the commitment score. The points were then summed to create the multiple offense score (top-coded at 12).[10] Offense type is a 4-cateogry nominal indicator for the type of crime committed: Violent crimes (including drug trafficking), property crimes, drug crimes, and other crimes. Dummy variables were created from this nominal measure with violent crimes serving as the reference category. Trial is a binary indicator of whether the offender was

[10] Note that the commitment score was constructed post hoc by the Commission and thus was not available to or considered by the sentencing judge. It is included as a proxy to measure the nature and number of offenses for those individuals sentenced after pleading guilty or being found guilty of multiple offenses, which are not otherwise accounted for, but which likely would be considered by the sentencing judge.

found guilty after a trial rather than entering a guilty plea. In addition, mandatory minimums can have pronounced effects on sentencing outcomes (Kautt and DeLone, 2006; Rehavi and Starr, 2014). Thus, we included mandatory minimum as a binary indicator identifying the 34 offense codes which carried a non-suspendable mandatory prison term.

Finally, we included several extralegal characteristics, the most important of which for our purposes is the race of the offender. Black is a dummy variable indicating whether the offender was African American or white (reference category). Male indicates the offender's gender (female is the reference category). Age is the age of the offender (in years) at the time of admission, and the quadratic term for age is included in the model to explicitly capture any potential nonlinear effects.

3.2 Analytic Strategy

We propose a method established for event counts that will allow researchers to properly estimate models for criminal sentencing outcomes.[11] From our perspective, criminal sentences are essentially counts of the number of months that an offender has been sentenced to prison. As the name implies, event counts measure the frequency of an event for a given observation period

[11] To address potential selection bias in this two-stage modeling, scholars sometimes incorporate a Heckman correction (e.g., Nobiling et al. 1998; Steffensmeier and Demuth 2001; Ulmer and Johnson 2004). However, Bushway, Johnson, and Slocum (2007) note that the execution of the Heckman correction in sentencing studies has been highly problematic. Among the most serious problems is the near ubiquitous failure to incorporate an exclusion restriction, which requires at least one variable that affects the selection process but not the substantive equation of interest. As a result, Bushway and colleagues caution that employing the Heckman correction may cause more harm than good in many instances. Since the Bushway et al. (2007) article was published, many scholars have considered but not reported Heckman corrected models (e.g., Doerner and Demuth, 2010; Johnson, Ulmer, and Kramer, 2008; Lieber and Johnson, 2008; Ulmer, Eisenstein, and Johnson, 2008). Were we to estimate an OLS model rather than the one we introduce here, we would also proceed without a Heckman correction because (1) we are unable to identify an exclusion restriction; and (2) the condition number for our independent variables, including interaction terms, is well above the suggested rule of thumb of 20 (see Bushway, Johnson, and Slocum, 2007: 168-69).

(Zorn, 1996). Modeling event counts can be problematic because they are non-negative integers bounded at zero, usually concentrated among a few small discrete values, and generally heteroskedastic with the variance increasing with the mean (Cameron and Trivedi, 2013). For example, the distribution of our sentencing measure is characterized by extreme positive skew, in which the mode and median are 0 months in prison (63% of the total cases), while the mean is a prison sentence of just over 16 months. The non-normality of sentencing outcomes can be seen graphically in Figure 1.

As Long (1997: 217) and others note, using ordinary least squares (OLS) regression to analyze untransformed event counts is inadvisable because OLS can cause "inefficient, inconsistent, and biased estimates." As a workaround, many criminal sentencing scholars have opted to use a log-transformed measure of sentence length for those who are incarcerated (e.g., Freiberger and Hilinski, 2013; Spohn and Cederblom, 1991; Spohn and DeLone, 2000; Steffensmeier and Demuth, 2000; Steffensmeier, Ulmer, and Kramer, 1998). This common practice has arisen largely because of the unique nature of sentencing dependent variables, most notably their highly skewed, non-negative, and intrinsically heteroscedastic distributions (Cameron and Trivedi, 2013).

Fischman and Schanzenbach (2012: fn.20) defend this practice in the sentencing context: "OLS regression with robust standard errors still provides consistent estimates, even when the error terms are not normally distributed." Yet, Santos, Silva, and Tenreyo (2006: 641) demonstrate that "in the presence of heteroskedasticity, estimates obtained using log-linearized models are severely biased, distorting the interpretation of the model." They show that this bias occurs because the expected value of a log-transformed variable depends on higher-order moments of its distribution. In other words, "if the errors are heteroskedastic, the transformed errors will be generally correlated with the covariates" (Santos, Silva, and Tenreyo, 2006: 653). Hilbe (2014: 17) also argues that "when the count response is logged and modeled using linear regression, its predicted values are nearly always distant from the actual or observed counts." Hilbe emphatically advises: "Reject the temptation to use linear regression to model a logged count" (Hilbe, 2014: 17, emphasis in original).[12]

When a count variable contains a high proportion of zeros and overdispersion such as the sentencing outcome in the current study, hurdle and zero-inflated models may be more appropriate than the Poisson or Negative Binomial distributions typically used to model event counts. Hurdle and zero-inflated models account for excess zeros by combining a binary model with a count model. For hurdle models, a logit (or probit) is used to predict observations that have zero counts, and a zero-truncated count model (e.g., the Negative Binomial) predicts the remaining non-zero cases (Cameron and Trivedi, 2013; Hilbe, 2014; King, 1988; Long and Freese, 2014; Mullahy, 1986; Zorn, 1996). In this two-stage process, we assume there is a threshold, or "hurdle," that must be surpassed in order to observe a positive count (e.g., time in prison). In theory, it is possible for any observation to cross this hurdle. In contrast, zero-inflated models assume that there are two distinct processes responsible for generating zeros in the data: One structural source, for which a positive count is never observed, and another source for which a positive count may or may not occur (Cameron and Trivedi, 2013; Lambert, 1992; Long and Freese, 2014; Zorn, 1996). Thus, the zero-inflated model simultaneously estimates a binary

[12] For our sentencing data we found evidence that an OLS model would indeed violate the assumption of homoscedasticity. Informal plots of the residuals versus fitted values and formal tests for heteroscedasticity such as Cameron and Trivedi's decomposition of IM-test, White's General Test, and the Breusch-Pagan / Cook-Weisberg Test detect heteroscedasticity indicate that the null hypothesis of constant error variance has been violated. Further, while logging the dependent variable does reduce heteroscedasticity in our data, we still find significant levels in diagnostic plots and tests of the transformed criminal sentence.

inflation equation that generates "excess" zeros, as well as a Poisson or Negative Binomial model for the remaining zero and non-zero counts.

In practice, the results from hurdle and zero-inflated models are usually very similar (e.g., see Zorn, 1996); yet, their underlying assumptions about the source of high zero counts should dictate which method is more appropriate. In our case, we have no reason to suspect that there is a structural source of zeros, in which a group of criminal offenders would never be sentenced to prison. For instance, even for the lowest level of offense seriousness and criminal history, the proportion of those incarcerated is .14 and .19, respectively.[13] Thus, a hurdle model seems more appropriate for our data, especially given that it maps onto existing theories of judicial decision-making in which judges are assumed to first decide whether to incarcerate an offender and then determine for how long (e.g., see Spohn and Cederblom, 1991; Spohn, 2009).

Formally, the Hurdle Regression Model using the Negative Binomial distribution (HRM-NB) is a combination of the logit model to predict 0s (Equation 1) and the modified zero-truncated negative binomial model to predict positive counts (Equation 2) (see Long and Freese, 2014, p. 520; 527-528):

$$\Pr(y_i = 0|\mathbf{x}_i) = \frac{\exp(x_i\gamma)}{1+\exp(x_i\gamma)} = \pi_i \quad [\text{Eq. } 1]$$

$$\Pr(y_i|\mathbf{x}_i) = (1-\pi_i)\left[\frac{\frac{\Gamma(y_i+\alpha^{-1})}{y_i!\,\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_i}\right)^{\alpha^{-1}}\left(\frac{\mu_i}{\alpha^{-1}+\mu_i}\right)^{y_i}}{1-(1-\alpha\mu_i)^{-1/\alpha}}\right] for\ y > 0 \quad [\text{Eq. } 2]$$

[13] Even the combination of minimal values for offense seriousness and criminal history leads to a non-trivial probability of incarceration, as the proportion of those sentenced to prison with these attributes is .034. A zero-inflated model might be appropriate in the sentencing context, however, where the data also included infractions for which active incarceration was never an option. We also note that Anderson, Kling, and Stith (1999) used a zero-inflated negative binomial model to examine inter-judge disparity before and after the federal sentencing guidelines became binding.

where γ is a set of covariates that determines the incarceration decision; $\Gamma(\cdot)$ is the gamma function, α is a dispersion parameter, and $\mu_i = \exp(\mathbf{x}_i\beta)$. In this two-equation model, 0 is conceptualized as the "hurdle" that must be passed before we can observe a positive count. Although these models can be estimated individually, "the two processes are conjoined using the log-likelihood…[which] is the log of the probability of y = 0 plus the log of y = 1 plus the log of y being a positive count" (Hilbe, 2014: 185).

As a brief demonstration that our hurdle model provides unbiased estimates of overdispersed count data with a high proportion of zeros, we conducted a series of Monte Carlo simulations testing different modeling strategies (for an accessible discussion of simulations, see Carsey and Hardin, 2014). To this end, we specified a data-generating process that mirrored our sentencing outcome and then regressed this variable on a single explanatory predictor using OLS and distributions for other count models.[14] We also tried to include a Heckman-corrected model using the logged sentencing outcome, which is a common approach in criminology, but this model repeatedly failed to estimate during the Monte Carlo simulations (possibly due to the absence of a required exclusion restriction). We present the results from these simulations in Table 2.

One advantage of the Monte Carlo simulations is that we know the true value of the population parameter: $\beta = 0.5$. As reported in Column 2 of Table 2, only the Hurdle Regression Model using the Negative Binomial (HRM-NB) and Zero-Inflated Negative Binomial (ZINB) models provide unbiased estimates of $\hat{\beta}$ (standard errors are listed in Column 3). The difference among the various models becomes even more apparent when we consider the percentage of

[14] The R replication code is available upon request.

relative bias resulting from each method in Column 4. The Ordinary Least Squares (OLS) model of the untransformed outcome provides wildly inaccurate estimates of the true effect (a relative bias of more than 1,000%), and while other count models come closer to the true estimate, only the models designed to account for zero-inflation and overdispersion produce trivial percentages of relative bias (< 0.1%). White's test confirms a significant degree of heteroscedasticity even for the log-transformed OLS model, ${\chi^2}_{(1)} = 25.2$, p < .001, which means that the OLS model of the log-transformed outcome will produce biased estimates because of the presence of heteroscedasticity. Column 5 contains the coverage probabilities from the various models, which are the proportions of the time that the confidence intervals actually contain the true value of interest. Again, only the HRM-NB and ZINB have acceptable coverage probabilities of .95; all of the remaining models have coverage probabilities of .10 or less. Finally, a comparative goodness-of-fit measure, the Akaike Information Criterion (AIC), is presented in the last column. In general, the lower the value of AIC, the better the model fit; thus, the HRM-NB or ZINB should be preferred over alternative specifications.

To summarize, the HRM-NB provides unbiased estimates of our highly skewed sentencing outcome, despite the overdispersion and the presence of a high proportion of real zeros in the data.[15] We now turn to the substantive findings that address our theoretical question

[15] For purposes of comparison, we specified three different counts models—the Poisson Regression Model (PRM), Negative Binomial Regression Model (NBRM), and the HRM-NB—to determine the best fit for our sentencing data. For the HRM-NB model, we used the same set of predictors for the incarceration decision as we did for the truncated sentencing count. First, we calculated mean predicted probabilities for each model, and then created a difference measure of the observed and predicted counts (Long and Freese, 2014). The ideal model would be one in which all plot points fall at 0, as this would indicate that our model perfectly predicted the observed data (i.e., observed – predicted = 0). The HRM-NB fit the data best, hovering closely around the reference line at zero. Second, because the NBRM reduces to the PRM when the overdispersion parameter, α, is equal to zero, a Likelihood Ratio (LR) test of the null hypothesis (H_0: α = 0) can be conducted. In our case, α > 0, and the resulting high value for the χ^2 statistic

whether, as the liberation hypothesis supposes, racial disparities are more likely to manifest at lower levels of crime seriousness or criminal history.

4. RESULTS

4.1 The Incarceration Decision

We begin by analyzing the effects of various case, offender, and extralegal characteristics on sentencing decisions. To this end, we regressed prison term on the set of case and offender characteristics described in the measures section, as well as two interaction terms that specifically test the liberation hypothesis: [16]

$$\begin{aligned} x_i\beta = {} & \beta_1(Offense\ Seriousness) + \beta_2(Commitment\ Score) + \beta_3(Drug\ Offense\ ) \\ & + \beta_4(Property\ Offense\ ) + \beta_5(Other\ Offense) + \beta_6(Trial) \\ & + +\beta_7(Mandatory\ Minimum) + \beta_8(Criminal\ History) + \beta_9(Male) \\ & + \beta_{10}(Age) + \beta_{11}(Age)^2 + \beta_{12}(Black) \\ & + \beta_{13}(Black \times Offense\ Seriousness) + \beta_{14}(Black \times Criminal\ History) \end{aligned}$$

We report the findings from the binary portion of the HRM-NB in Table 3, which shows the factors predicting whether an offender will be incarcerated. To help provide a meaningful interpretation of these nonlinear results, we also present average marginal effects (AMEs) in

---

led us to reject the null hypothesis and conclude that the NBRM should be preferred over the PRM. Third, Greene (1994) proposes using a Vuong test (V) for non-nested models like the NBRM and the HRM-NB. If V > 1.96, the first model is preferred; if V < -1.96, the second model provides a better fit. Using guidance provided by Long and Freese (2014), we computed V for the PRM vs. HRM-NB, as well as the NBRM vs. HRM-NB. The results of the Vuong test strongly support the HRM-NB over alternative count models such as the PRM and the NBRM, as V is well below the specified cutoff of -1.96.

[16] We also have data on the judge responsible for the sentencing decision for more than 99% of the cases in our data, which is important as we would expect sentencing decisions to be clustered by judge. Moulton (1990) demonstrates that failure to properly account for clustering can lead to massive underestimation of standard errors and flawed hypothesis tests.

Table 3 and graphically in Figure 2, which provide metrics of discrete or instantaneous rates of change in predictions for different values of interest.[17] The main effects are generally in keeping with recent studies from guidelines jurisdictions and show that legal characteristics like offense severity, prior record, and mandatory minimum offenses are among the strongest predictors of the in/out decision. For example, a one-level increase in offense severity (e.g., moving from a Serious Misdemeanor to a Class F Felony) increases the likelihood of being incarcerated by .09. Thus, moving from the least to most serious offense increases the predicted probability of incarceration by .36 (given that offense seriousness is measured on a 5-point scale). Each one-category step up in criminal history increases the likelihood of incarceration by .13, and moving from the lowest to highest score on this 5-point measure amounts to change in predicted probabilities of .52. The mode of disposition exerts a substantial impact on the likelihood of a prison sentence: the AME indicates that the difference in predicted probabilities is .41 for a trial conviction versus a guilty plea. Consistent with the prior literature and expectations, most of the extralegal characteristics exert a modest but statistically significant influence. The difference in incarceration rates for males relative to females is .08, while the difference is .06 for black offenders. However, age (and its squared term) does not appear to influence the outcome.

[17] For a dummy variable, the AME is the mean of differences in predictions for each observation (leaving all other values unchanged in the data) when moving from 0 to 1 for that variable. For example, the marginal effect of race for a single observation is the difference in the predicted number of months sentenced to prison assuming that the offender's race was first coded as 'white' and then as 'black'. To obtain the AME, we simply take the mean of these individual marginal effects, which allows us to compare the effect of race for two hypothetical populations—one all black and one all white—on criminal sentencing decisions (for more information about marginal effects, see Williams, 2012). For a continuous variable, the AME is the mean of instantaneous rates of change—that is, the mean of the slopes—at each value of the variable over all observations (leaving the rest of the data unchanged). Thus, the AME for a continuous variable provides a good approximation for the amount of change in Y given a 1-unit change in X_i.

For the interaction terms, only black x criminal history is statistically significant and consistent with the liberation hypothesis. The results show that offense seriousness does not significantly moderate the effects of race for the decision to incarcerate. Thus, contrary to the liberation hypothesis, the effect of race is not more pronounced at lower levels of offense seriousness where judges were theorized to have more freedom to allow extralegal characteristics like race to influence their decisions. Figure 3 provides a plot of the marginal effects of being black relative to white at different values of criminal history. As predicted, blacks are more likely to be sent to prison compared to whites only at lower levels of prior record; this incarceration disparity decreases as criminal history increases. The difference in the probability of incarceration for blacks with no or minimal criminal history is .07 relative to whites, while at the highest levels of criminal history blacks and whites have the same likelihood of avoiding prison. In sum, the effect of being black is mitigated as criminal history increases in keeping with the liberation hypothesis.

4.2 The Sentence Length Decision

As reported in Table 4, legal characteristics are the strongest drivers of the prison length determination, and there is also a strong trial penalty. Offense seriousness, current commitment score, trial, and mandatory minimum offenses all increase the number of months imposed, while being sentenced for property, drug, and other offenses compared to violent offenses are all associated with shorter prison terms. For the offender characteristics, criminal history and the age measures are not significant, while males and blacks are sentenced differently than females and whites.[18] The AMEs reported in Column 2 of Table 4 and graphically in Figure 4 reveal the

---

[18] We also specified a generalized negative binomial model using the gnbreg command in Stata which allowed us to explicitly model the predictors that contribute to ovedispersion in the count response. We discovered that many of the case characteristics—but not offender attributes—

most substantial effects on sentence length include the legal characteristics of offense severity, mandatory minimum offense, and the trial disposition indicator. Each severity level increases the average prison term by over two years while being convicted of a mandatory minimum offense is associated with an additional 41 months. Being found guilty after trial rather than a plea is particularly salient, as it adds nearly 5 years to the length of the average prison term. Among the extra-legal offender characteristics, males receive about additional 3 months compared to females, and blacks receive on average 5 fewer months compared to whites.

For the interaction terms, offense seriousness moderates race in a manner somewhat consistent with the liberation hypothesis, as we find a significant black x offense seriousness interaction. Figure 5 plots the marginal effect of being black at different levels of offense seriousness. This figure shows that black offenders are more likely to receive additional prison terms for less serious offenses relative to whites, although the effect size is relatively small. This difference in sentences equates to an additional 2 months in prison for serious misdemeanors, Class F Felonies, and Class E Felonies; the marginal effect reverses directions for Class D Felonies but is not statistically significant. For the most serious felonies, the marginal effects of show a net advantage of approximately 1 year less for black offenders. It is worth noting that the standard errors are much larger for this marginal effect (i.e., roughly 10 times the size of the standard errors for the lowest severity category). There is a notable curvilinear pattern in which the effect of being black slightly increases across low and mid-level severity offenses before dropping drastically for the most severe offenses.

Given that criminal history did not significantly predict positive prison sentences, we conducted supplemental analyses in which we dummy coded criminal history rather than treating

account for overdispersion: offense seriousness, type of offense (drug, property, and other), trial, and mandatory minimum are significant predictors.

it as an ordinal measure. Several scholars have reported that this alternative specification has increased the explained variance of their sentencing models in guidelines jurisdictions (e.g., Bushway and Piehl, 2001; Engen and Gainey, 2000; Mustard, 2001), and they have also found that the effects sizes of extralegal characteristics like race actually declined as a result (Bushway and Piehl, 2001). However, our dummy coding of criminal history did not increase model fit, and the proportion of explained variance in our sentencing outcome remained virtually unchanged for these non-guidelines data (i.e., Nagelkerke R^2 for both models was .71). In addition, the dummy coding of criminal history increased the number of parameters in the model from 2 to 8, thereby also increasing the likelihood of committing a Type I error. Differences in the main effect of history were minor and depended upon the excluded category; in most cases, the black x criminal history interactions were not statistically significant leading us to conclude that there was little added benefit to the alternative model.[19]

5. DISCUSSION

The liberation hypothesis posits that for the most serious criminal offenses and the most seasoned repeat criminal offenders, judges will feel little choice but to impose severe punishment regardless of extralegal factors like race. Yet, in more ambiguous contexts, judges will be "liberated" from the constraints of extreme severity and criminality; in these instances the door opens for extralegal characteristics such as race to influence sentencing decisions. We tested this

[19] For the main effects of history in the dummy-coded model, the no-history group appeared to receive longer average sentences than some of the other categories. This seems counterintuitive, but most offenders with no criminal history would not be expected to be imprisoned in the first place; for those first offenders who were sent to prison, it is possible that some aspect of the case or offender that led to the exceptional disposition of prison (e.g., harm to victim, the demeanor or attitude of the defendant, etc.), also resulted in a comparatively punitive sentence length. Where differences were statistically significant, they were modest at best. Results are available upon request from the first author.

theory in a jurisdiction without the constraints of sentencing guidelines; the large grants of discretion in South Carolina make it an ideal location for examining the hypothesis.

We found strong support for the liberation hypothesis when considering how race interacts with criminal history for the incarceration decision. These results are consistent with the idea that when offenders have accrued a substantial record of past convictions, their criminal history has a constraining effect that neutralizes the influence of race: judicial decision makers are equally likely to impose a prison term on offenders with extensive records, regardless of their race. However, for the many offenders who do not have extensive records, presumably a whole panoply of factors can influence the in/out decisions of courtroom workgroups. Among these low-history offenders, race effects become significant and fairly strong, with blacks being 5%-7% more likely to be incarcerated compared to whites. Once the decision-making process moves to the sentence length determination, the liberation mechanism appeared to operate differently. For instance, we did not find a statistically significant race x criminal history interaction; instead we found some support for the hypothesis with a significant race x severity interaction. Low severity black offenders received slightly longer sentences than whites; yet, high severity black offenders appear to have received shorter average sentences than whites. The standard errors and resulting confidence intervals are quite large for this latter effect leading us to consider these results with some caution. It is possible that with these high severity offenses, unmeasured case-specific facts about the victim and the nature of the conduct render the race effects unstable.

Consequently, our study adds to the body of studies which find mixed or equivocal support for the liberation hypothesis. On one hand, our results harmonize with Spohn and Cederblom (1991) who found support for the hypothesis for the disposition decision but not the length duration. However, even this pattern is not universal. Spohn and DeLone (2000) found support for the liberation hypothesis in some jurisdictions but not others, and Warren et al. (2012) report a complex set of findings in which support for the hypothesis varies among various offender/case/outcome combinations. It is possible that the liberation hypothesis is simply too parsimonious to accurately account for many of the individual, group, and contextual processes at play in sentencing decisions. One of many possibilities here involves the potentially unseen impact of the victim. The majority of violent crime is intra-racial, and one theory of bias in the criminal justice system is that court decision-makers undervalue black victims (see Baldus, Pulaski, and Woodworth, 1983; Blumstein, 1993). Since black victims are most likely to be victimized by black offenders, if courtroom actors do sentence crimes with black victims less punitively, this practice could confound the race x severity interaction and actually result in a black advantage at high severity levels that encompass violent crimes against the person. Perhaps the most effective line of inquiry into these different processes would be qualitative research among courtroom workgroup members—a call that has been renewed by many in recent years (see Baumer, 2013; Ulmer, 2012).

Interestingly, the main effect of criminal history was not a significant predictor of the sentence length determination. This is a particularly noteworthy finding coming from a non-guidelines state, since guidelines jurisdictions build in prior record as one of the two determinants of both the incarceration and sentence length decision. Sentencing scholars have noted that in this way some guidelines may actually be building racial disparities into the formal sentencing structure, as black offenders typically have more significant criminal records than whites (Tonry 1995, 1996). Frase (2009), for example, demonstrated that in Minnesota a full two-thirds of the race disparity in the guidelines recommendations of presumptive prison sentences was attributable to criminal history. The South Carolina judges in the current study, unconstrained by sentencing guidelines, did consider prior record to be highly relevant for the decision to incarcerate—in fact, it was the third strongest predictor following mandatory minimum and trial disposition. However, criminal history did not influence sentence length for those sent to prison. If the typical sentencing guidelines grid had been in effect in South Carolina, offenders with longer prior records would presumably have been recommended for much longer sentences than these judges deemed appropriate. While this criminal history finding seems counterintuitive, it is not without precedent. For example, in Levin's (1977: 102) classic study, he observed that "in Minneapolis defendants with a prior record receive much more severe sentences than those without one. This pattern does not occur in Pittsburgh sentencing." Thus, while it is highly likely that criminal history has some relevance in most jurisdictions, the significance of prior record may well be a component of the legal culture of a jurisdiction that varies according to outcome and location.

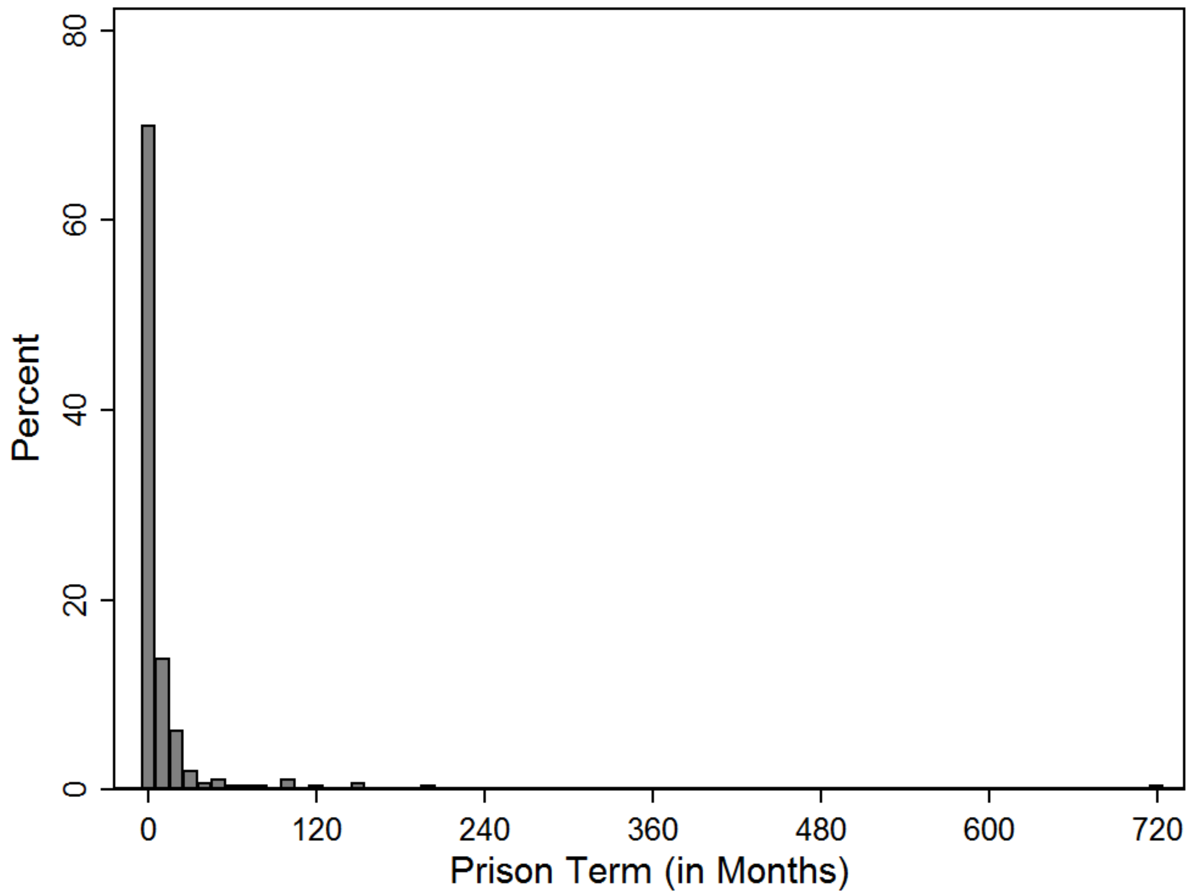
6. CONCLUSION

The current research offers several contributions to the literature. Given the indeterminate nature of South Carolina sentencing law and its lack of guidelines to constrain judicial decision-making, these data provide a unique opportunity to test the liberation hypothesis since judicial discretion, and thus the opportunity for extralegal disparities to manifest, should be present to a greater extent than that found in guidelines jurisdictions where most recent sentencing research has been conducted. However, there are several notable limitations to our research. Our data are from only one year and one jurisdiction. Like much of the sentencing research, we lack potentially important controls such as whether an offender had been detained prior to trial, his or her socioeconomic status, whether they had dependent children, and information related to the victim. And since the jurisdiction is in the South and in a state with a comparatively high black population, it is unclear whether similar results would be found in non-guidelines jurisdictions in other regions of the country.

The results from our analyses provide some support for the liberation hypothesis. First, for the in/out decision, the main effect for race was statistically significant, with blacks being about 6% more likely to be incarcerated holding the other relevant variables constant. This black penalty did not vary by offense seriousness as proposed by the liberation hypothesis, but an even greater effect of race did exist at lower levels of prior record. This finding was consistent with the hypothesis that with extensive offenders, judges will feel little choice but to incarcerate, while at lower criminal history levels the extralegal effect of race becomes more prevalent. With the sentence length decision, the moderating factors differed: no race differences were found for criminal history, but blacks were more likely to be sentenced to longer terms of incarceration when convicted for less severe offenses (yet received shorter sentences for the most severe category of crimes).

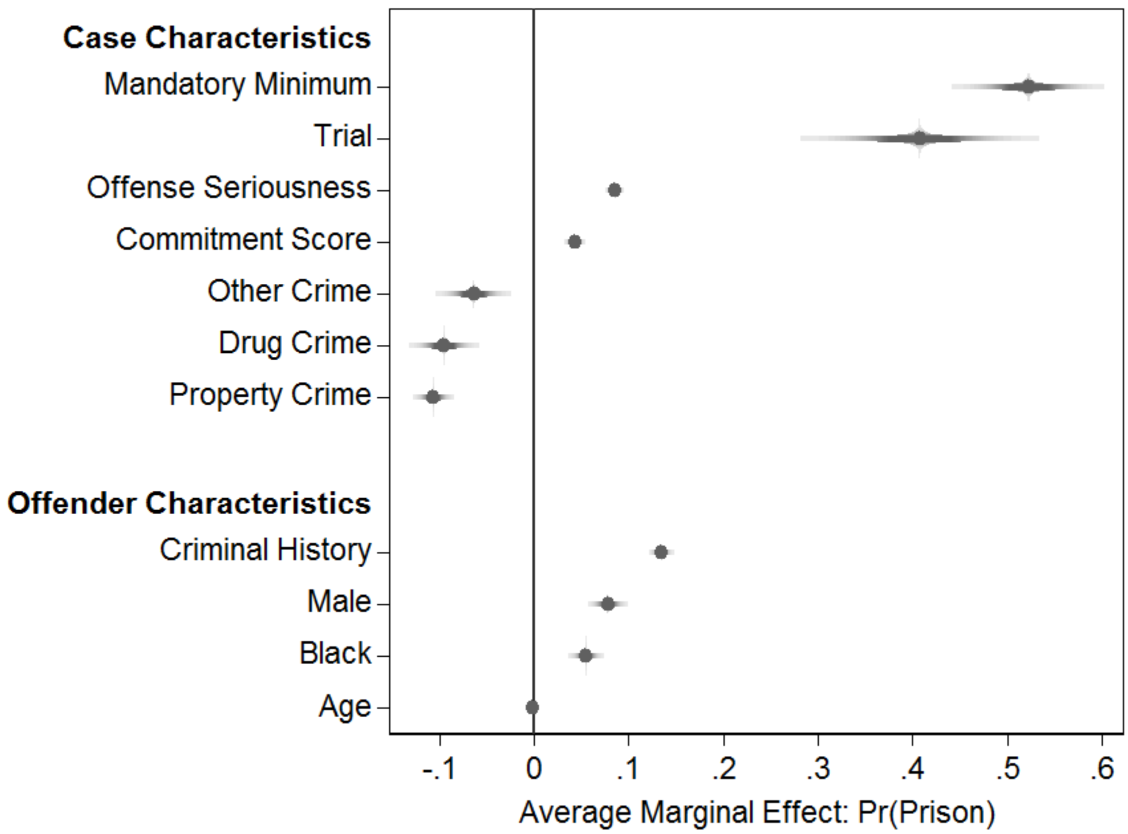
To test for the liberation hypothesis, we employed a hurdle count model which aptly fit the positively skewed imprisonment decision frequently found in sentencing data. While similar count models have been used in other areas of criminology and criminal justice (e.g., see MacDonald and Lattimore, 2010), researchers have yet to embrace them in sentencing studies. We demonstrated through a series of Monte Carlo simulations that the hurdle model returned unbiased estimates compared to alternative modeling approaches. As sentencing scholars continue to utilize advancements in statistical modeling strategies (see, e.g., Johnson, 2012; MacDonald and Lattimore, 2010) researchers may wish to consider count models given the nature and distribution of sentencing outcomes.

Figure 1. Distribution of Event Counts, Expected Minimum Prison Term (in Months)



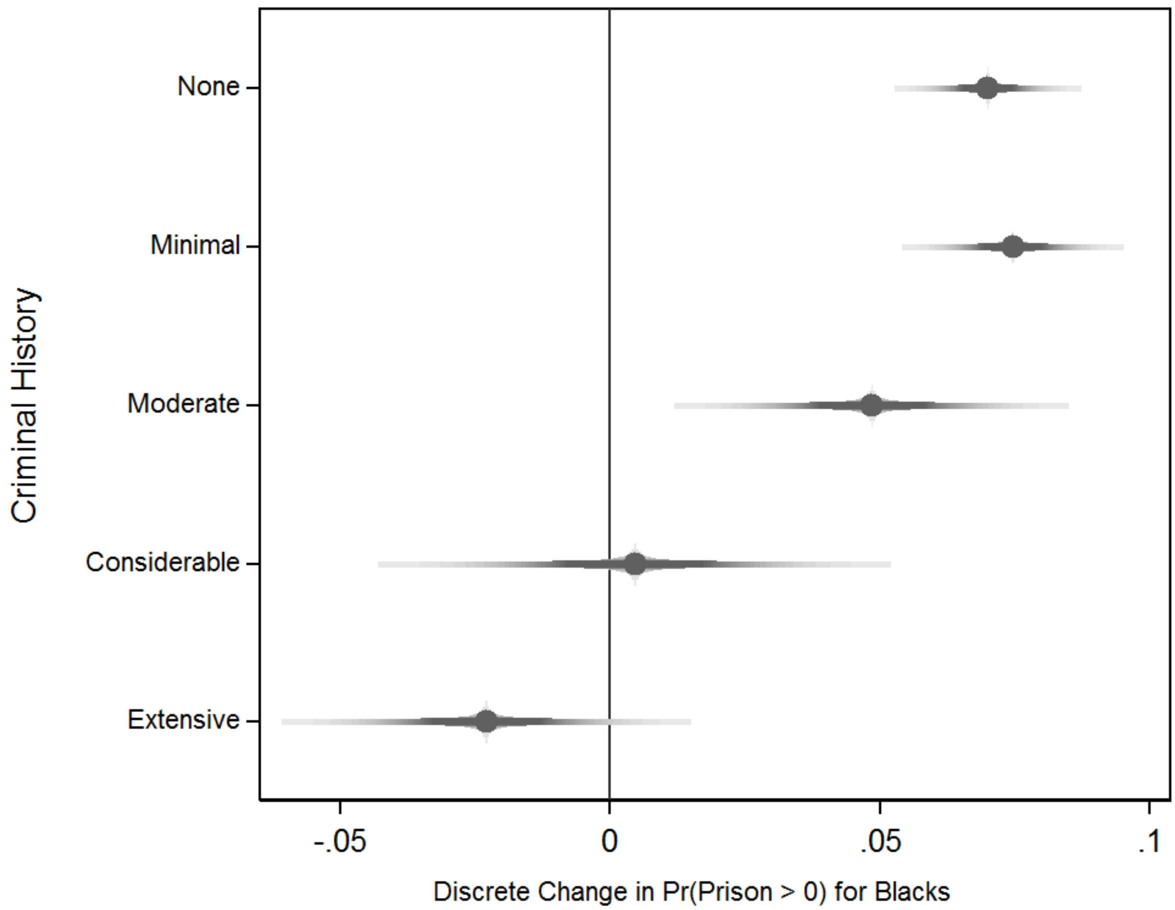
Notes: N = 17,671. Expected minimum prison terms range from 0 to 720 months; M = 16.2, SD = 64.4; Mdn = 0; Mode = 0.

Figure 2. Average Marginal Effects of Case and Offender Characteristics on the Probability of Being Incarcerated.



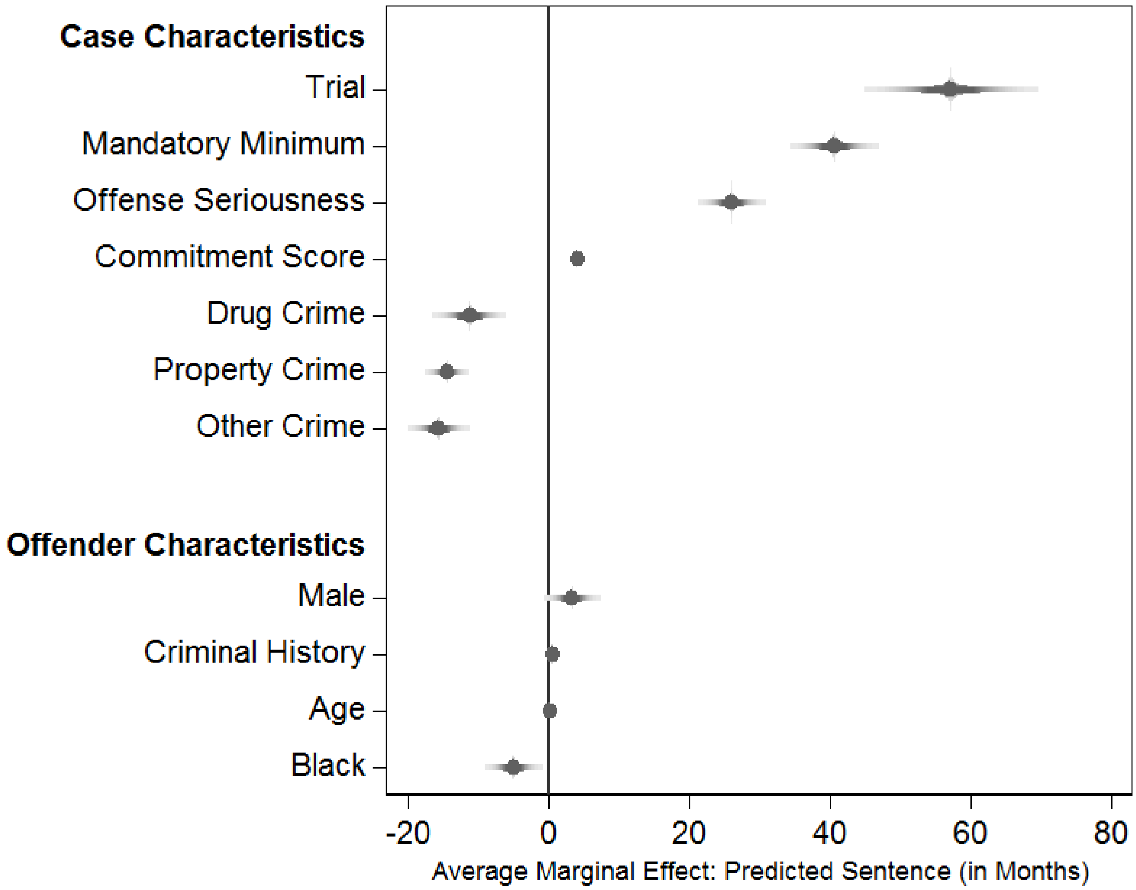
Notes: Average Marginal Effects (AMEs) are calculated from the logistic regression equation of the hurdle regression model with bootstrapped standard errors (1,000 replications clustered by sentencing judge, 51 in total). For dummy variables, the AME is the mean of differences in predictions for each observation when moving from 0 to 1, leaving the rest of the data unchanged. For continuous predictors, the AME is the mean of instantaneous rates of change at each value for every observation, leaving the rest of the data unchanged. Confidence intervals surround point estimates. N = 17,671.

Figure 3. Marginal Effect of Being Black on the Likelihood of Incarceration (Black – White)



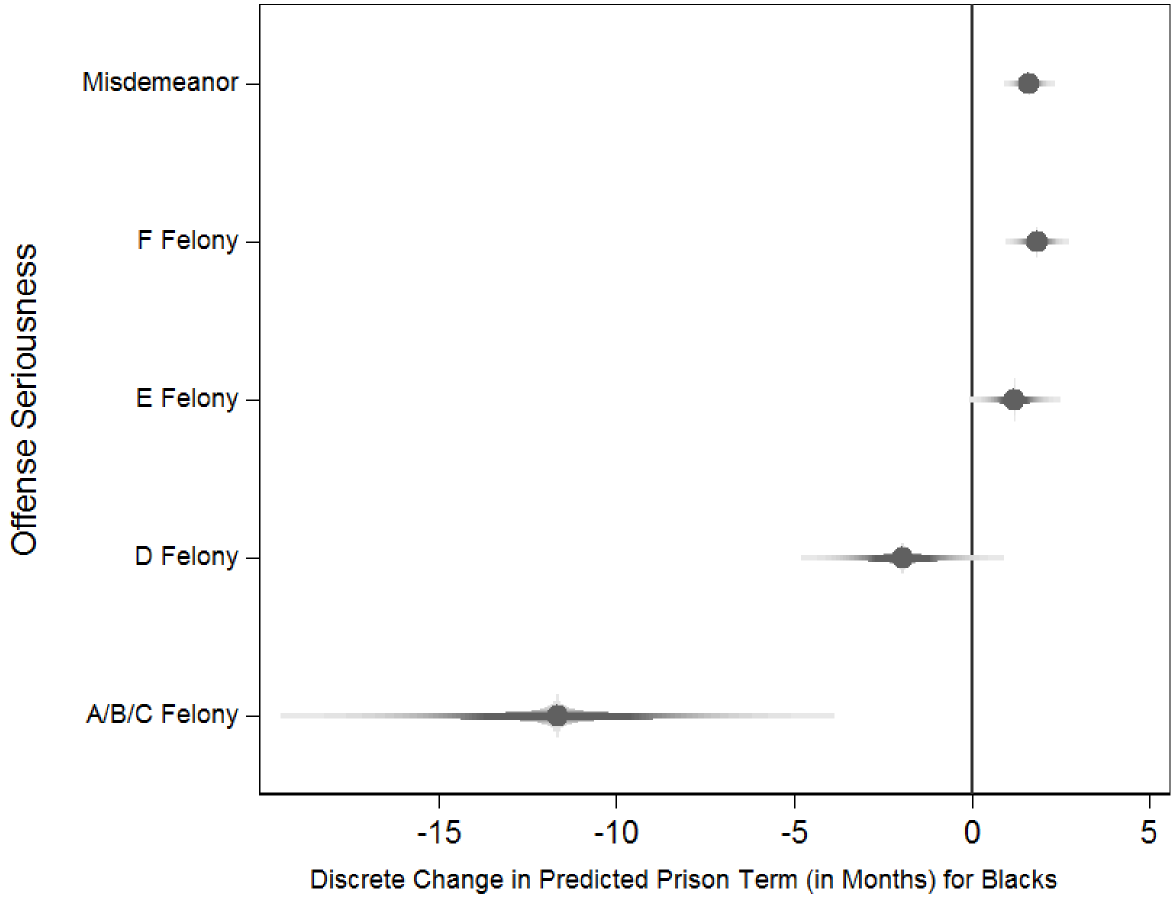
Notes: Marginal effects at representative values (MERs) are derived from the logistic regression equation of the hurdle regression model with bootstrapped standard errors (1,000 replications clustered by sentencing judge, 51 in total). MERs compare the predicted probability of being incarcerated for black vs. white offenders. Positive plot points indicate that a black offender is more likely to be incarcerated relative to a white offender depending upon his or her criminal history. Error bars are the 95% confidence intervals for the estimated marginal effects. N = 17,671.

Figure 4. Average Marginal Effects of Case and Offender Characteristics on Predicted Prison Sentence (in Months) for those Incarcerated.



Notes: Average Marginal Effects (AMEs) are calculated from the zero-truncated negative binomial equation of the hurdle regression model with bootstrapped standard errors (1,000 replications clustered by sentencing judge, 51 in total). For dummy variables, the AME is the mean of differences in predictions for each observation when moving from 0 to 1, leaving the rest of the data unchanged. For continuous predictors, the AME is the mean of instantaneous rates of change at each value for every observation, leaving the rest of the data unchanged. Confidence intervals surround point estimates. N = 17,671.

Figure 5. Marginal Effect of Being Black on the Predicted Prison Term (in Months)



Notes: Marginal effects at representative values (MERs) are derived from the count equation of the ZINB model. MERs compare the predicted count (number of months in Prison) for black vs. white offenders at different values of offense seriousness. MERs compare the predicted prison term for incarcerated black vs. white offenders. Positive plot points indicate that a black offender receives a longer prison term than a white offender depending upon the seriousness of the offense. Predicted counts are listed above point estimates, and error bars are the 95% confidence intervals for the estimated marginal effects. N = 17,671.

Table 1. Description of Variables

Variable	Description	Code	Summary Statistics	
Dependent Variable			%	n
Prison Term	Expected minimum # of months sentenced (0 to 720 months)		Mean = 16.2 Std. Dev. = 64.4	
Case Characteristics				
Offense Seriousness	5-level ordinal score From the S.C. Crime Classification Scheme	1 = Misdemeanor	15.3%	2,694
		2 = Felony (Class F)	46.0%	8,136
		3 = Felony (Class E)	19.6%	3,455
		4 = Felony (Class D)	11.0%	1,942
		5 = Felony (Class A, B, C, or Unclassified)	8.2%	1,444
Commitment Score	12-level ordinal measure Number of commitment offenses	1 = Less Serious 12 = Very Serious	Mean = 1.9 Std. Dev. = 1.8	
Offense Type	4-category nominal indicator of the classification of crime committed (Violent offenses are the reference category)	1 = Violent	34.3%	6,065
		2 = Drug	14.5%	2,561
		3 = Property	33.2%	5,858
		4 = Other	18.0%	3,187
Mandatory Minimum	Minimum prison sentence mandated	1 = Yes	5.1%	897
		0 = No	94.9%	16,746
Trial	Found guilty after trial	1 = Guilty after Trial	1.5%	258
		0 = Guilty Plea	98.5%	17,413
Offender Characteristics				
Criminal History	5-level ordinal score Derived from the number and severity of prior offenses	1 = None	36.6%	6,460
		2 = Minimal	32.9%	5,806
		3 = Moderate	17.4%	3,080
		4 = Considerable	5.9%	1,048
		5 = Extensive	7.2%	1,277
Black	Race	1 = Black	62.0%	10,950
		0 = White	38.0%	6,721
Male	Gender	1 = Female	16.6%	2,929
		0 = Male	83.4%	14,742
Age	Age (15 to 81 years old)		Mean = 31.3 Std. Dev. = 10.1	

Notes: Percentages may not sum to 100% due to rounding errors.

Table 2. Monte Carlo Simulations of Methods for Analyzing Overdispersed Count Data with Zero-Inflation

Model	$\boldsymbol{\hat{\beta}}$ $\boldsymbol{(True = 0.50)}$	$\boldsymbol{\widehat{SE}}$	**Relative Bias (%)** $\boldsymbol{= \frac{(\hat{\beta} - \beta)}{\beta}(100)}$	**Coverage Probability**	**AIC**
Ordinary Least Squares (OLS)	5.51	0.84	1,001.69%	.00	11,481.96
OLS Log-Linear Model (ln(y) if y > 0)*	0.40	0.02	--	--	1,125.26
Poisson (PRM)	0.37	0.00	-25.40%	.01	67,971.73
Negative Binomial (NBRM)	0.38	0.04	-23.73%	.05	4,214.81
Hurdle Poisson (HRM-P)	0.47	0.00	-5.64%	.10	22,895.78
Zero-Inflated Poisson (ZIP)	0.47	0.00	-5.62%	.10	22,889.29
Hurdle Negative Binomial (HRM-NB)	**0.50**	**0.03**	**-0.07%**	**.95**	**4,042.28**
Zero-Inflated Negative Binomial (ZINB)	0.50	0.03	-0.06%	.94	4,032.86

Notes: Simulations were conducted in R with 10,000 repetitions of N = 1,000 and a seed value of 8675309. The coverage probability reveals the proportion of estimated confidence intervals for the simulated samples which contain the true population parameter (Carsey and Hardin, 2014). Akaike Information Criterion (AIC) is a comparative fit statistic.

* The log-linear model for positive values would be combined with a logit/probit model to predict zero counts. AIC for the log-linear model is only a partial fit statistic (i.e., when y > 0). Relative bias and coverage probabilities were not calculated for the log-linear OLS model because the estimates are not on the raw-scale of y.

Table 3. Hurdle Model: Logit Estimates for the Incarceration Decision

	Incarceration Decision	Average Marginal Effect
Case Characteristics		Pr(Prison)
Offense Seriousness	0.68*** (0.06)	0.09*** (0.00)
Commitment Score	0.32*** (0.03)	0.04*** (0.00)
Drug Offense	-0.67*** (0.10)	-0.09*** (0.01)
Property Offense	-0.76*** (0.06)	-0.11*** (0.01)
Other Offense	-0.44*** (0.10)	-0.06*** (0.02)
Trial	2.69*** (0.38)	0.41*** (0.05)
Mandatory Minimum	3.72*** (0.43)	0.52*** (0.03)
Offender Characteristics		
Criminal History	1.15*** (0.05)	0.13*** (0.01)
Male	0.60*** (0.07)	0.08*** (0.01)
Age	-0.01 (0.01)	-0.00*** (0.00)
Age^2	-0.00 (0.00)	
Black	1.14*** (0.17)	0.06*** (0.01)
Liberation Hypothesis		
Black x Offense Seriousness	**-0.07 (0.04)**	
Black x Criminal History	**-0.24*** (0.04)**	
Constant	-5.56*** (0.31)	
Cragg and Uhler's R^2	0.51	
AIC_H	0.85	
Non-zero Observations (n)	6,611	
N	17,671	

Notes: Judges' sentencing decisions (prison term in months) modeled using the Hurdle Regression Model (HRM) with bootstrapped standard errors (1,000 replications clustered by sentencing judge; 51 in total). For dummy variables, the average marginal effect (AME) is the mean of differences in predictions for each observation when moving from 0 to 1, leaving the rest of the data unchanged. For continuous predictors, the AME is the mean of instantaneous rates of change at each value for every observation, leaving the rest of the data unchanged. Standard errors shown in parentheses. AIC_H is an enhanced Akaike Information Criterion comparative fit test (Hilbe, 2014). *** p < .001; ** p < .01; * p < .05.

Table 4. Hurdle Model: Zero-Truncated Negative Binomial Estimates for the Count Equation

	Prison Tern	Average Marginal Effect
Case Characteristics		in Months
Offense Seriousness	0.69*** (0.02)	26.08*** (1.64)
Commitment Score	0.10*** (0.01)	4.07*** (0.32)
Drug Offense	-0.27*** (0.05)	-11.16*** (2.06)
Property Offense	-0.36*** (0.03)	-14.38*** (1.23)
Other Offense	-0.40*** (0.04)	-15.61*** (1.71)
Trial	0.95*** (0.05)	57.18*** (4.80)
Mandatory Minimum	0.94*** (0.04)	40.63*** (2.43)
Offender Characteristics		
Criminal History	0.01 (0.02)	0.56 (0.39)
Male	0.08* (0.04)	3.34* (1.56)
Age	-0.00 (0.01)	0.26** (0.09)
Age^2	0.00 (0.00)	
Black	0.39*** (0.09)	-4.95** (1.59)
Liberation Hypothesis		
Black x Offense Seriousness	**-0.11*** (0.02)**	
Black x Criminal History	**0.00 (0.02)**	
Constant	0.26* (0.17)	
Log α	-0.59*** (0.03)	
Cragg and Uhler's R^2	0.72	
AIC_H	7.63	
N	6,611	

Notes: Judges' sentencing decisions (prison term in months) modeled using the Hurdle Regression Model (HRM) with bootstrapped standard errors (1,000 replications clustered by sentencing judge; 51 in total). For dummy variables, the average marginal effect (AME) is the mean of differences in predictions for each observation when moving from 0 to 1, leaving the rest of the data unchanged. For continuous predictors, the AME is the mean of instantaneous rates of change at each value for every observation, leaving the rest of the data unchanged. Standard errors shown in parentheses. AIC_H is an enhanced Akaike Information Criterion comparative fit test (Hilbe, 2014). *** p < .001; ** p < .01; * p < .05.

REFERENCES

Albonetti, C. A. (1991). An integration of theories to explain judicial discretion. Social Problems 38:247-66.

Anderson, J. M., Kling, J. R., & Stith, K. (1999). Measuring interjudge sentencing disparity: Before and after the federal sentencing guidelines.

Baldus, D. C., Pulaski, C., and Woodworth, G. (1983). Comparative review of death sentences: An empirical study of the Georgia experience. Journal of Criminal Law and Criminology, 661-753.

Baldus, D. C., Woodworth, G., and Pulaski, C. A. (1990). Equal justice and the death penalty: A legal and empirical analysis. Upne.

Ball, J. D. (2006). Is it a prosecutor's world? Determinants of count bargaining decisions. Journal of contemporary criminal justice, 22(3), 241-260.

Barkan, S. E., and Cohn, S. F. (1998). Racial prejudice and support by whites for police use of force: A research note. Justice Quarterly, 15(4), 743-753.

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

Blumer, H. (1969). Symbolic interactionism: Perspective and method. Prentice-Hall: Englewood Cliffs, N.J.

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

Blumstein, A. (1982). On the racial disproportionality of United States' prison populations. J. Crim. l. and Criminology, 73, 1259.

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

Bushway, S. D., and Piehl. A. M. (2001). Judging judicial discretion: Legal factors and racial discrimination in sentencing. Law and Society Review 35:733-64.

Cameron, A. C. and Trivedi, P. K. (2013). Regression of Count Data, 2nd Edition. New York, NY: Cambridge University Press.

Carsey, Thomas M., and Harden, Jeffrey J. (2014). Monte Carlo Simulation and Resampling Methods for Social Science. Thousand Oaks, CA: Sage Publications.

Chen, E. Y. (2008). The Liberation Hypothesis and Racial and Ethnic Disparities in the Application of California's Three Strikes Law. Journal of Ethnicity in Criminal Justice, 6(2), 83-102.

Chiricos, T. G. and W. D. Bales (1991). Unemployment and punishment: An empirical assessment. Criminology 29:701-24.

Chiricos, T. G., and C. Crawford. (1995). Race and imprisonment: A contextual Assessment of the evidence. In Ethnicity, Race, and Crime: Perspectives Across Time and Space, edited by D. F. Hawkins. Albany: State University of New York Press.

Doerner, J. K., & Demuth, S. (2010). The independent and joint effects of race/ethnicity, gender, and age on sentencing outcomes in US federal courts. Justice Quarterly, 27(1), 1-27.

Eisenstein, J., Flemming, R. B., & Nardulli, P. F. (1988). The Contours of Justice: Communities and their Courts. Boston, MA: Little, Brown.

Engen, R. L. (2009). Assessing determinate and presumptive sentencing—Making research relevant. Criminology and Public Policy 8(2):323-36.

Engen, R. L., and Gainey, R. R. (2000). Modeling the effects of legally relevant and extralegal factors under sentencing guidelines: the rules have changed. Criminology, 38(4), 1207-1230.

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

Frase, R. S. (2009). What Explains Persistent Racial Disproportionality in Minnesota's Prison and Jail Populations?. Crime and Justice, 38(1), 201-280.

Freiburger, T. L., and Hilinski, C. M. (2013). An examination of the interactions of race and gender on sentencing decisions using a trichotomous dependent variable. Crime and Delinquency, 59(1), 59-86.

Gertz, M. G. and Price, A. C. (1985). Variables influencing sentencing severity: Intercourt differences in Connecticut. Journal of Criminal Justice 13(2):131-39.

Grattet, R., and Lin, J. (2014). Supervision Intensity and Parole Outcomes: A Competing Risks Approach to Criminal and Technical Parole Violations. Justice Quarterly, (ahead-of-print), 1-19.

Greene, W. H. (1994). Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models (March 1994). NYU Working Paper No. EC-94-10. Available at SSRN: http://ssrn.com/abstract=1293115

Guevara, L., Boyd, L. M., Taylor, A. P., and Brown, R. A. (2011). Racial disparities in juvenile court outcomes: A test of the liberation hypothesis. Journal of Ethnicity in Criminal Justice, 9(3), 200-217.

Hilbe, J. M. (2014). Modeling Count Data. New York: Cambridge University Press.

Johnson, B. D., Ulmer, J. T., & Kramer, J. H. (2008). The social context of guidelines circumvention: The case of federal district courts. Criminology, 46(3), 737-783.

Johnson, B. D. (2012). Cross-classified multilevel models: an application to the criminal case processing of indicted terrorists. Journal of Quantitative Criminology, 28(1), 163-189.

Kalven, H., Zeisel, H., Callahan, T., and Ennis, P. (1966). The American Jury (p. 498). Boston: Little, Brown.

Kautt, P. M., & Delone, M. A. (2006). Sentencing outcomes under competing but coexisting sentencing interventions: Untying the Gordian knot. Criminal Justice Review, 31(2), 105-131.

King, G. (1988). Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for The Exponential Poisson Regression Model. American Journal of Political Science 32: 838-863.

Kramer, J. and Ulmer, J. (2009). Sentencing Guidelines: Lessons from Pennsylvania. Lynne Rienner: Boulder, CO.

Kutateladze, B. L., Andiloro, N. R., Johnson, B. D., and Spohn, C. C. (2014). Cumulative disadvantage: Examining racial and ethnic disparity in prosecution and sentencing. Criminology, 52(3), 514-551.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics, 34(1), 1-14.

Leiber, M. J., and Blowers, A. N. (2003). Race and misdemeanor sentencing. Criminal Justice Policy Review, 14(4), 464-485.

Leiber, M. J., & Johnson, J. D. (2008). Being Young and Black What Are Their Effects on Juvenile Justice Decision Making?. Crime & Delinquency, 54(4), 560-581.

Levin, M. A. (1977). Urban Politics and the Criminal Courts. University of Chicago Press.

Lin, J., Grattet, R., and Petersilia, J. (2012). Justice by other means: Venue sorting in parole revocation. Law and Policy, 34(4), 349-372.

Long, J. S. (1997). Regression models for categorical and limited dependent variables. Advanced quantitative techniques in the social sciences, 7.

Long, J. S., and Freese, J. (2014). Regression models for categorical dependent variables using Stata. Stata press.

MacDonald, J. M., and Lattimore, P. K. (2010). Count models in criminology. In Handbook of quantitative criminology (pp. 683-698). Springer New York.

McAninch, W. S., Fairey, F. W., and Coggiola, L. M. (2007). The Criminal Law of South Carolina (3rd ed.). Columbia, SC: South Carolina Bar.

Mitchell, O. (2005). A meta-analysis of race and sentencing research: Explaining the inconsistencies. Journal of Quantitative Criminology 21 (4):439-66.

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. The Review of Economics and Statistics, 334-338.

Mullahy, J. (1986). Specification and testing of some modified count data models. Journal of econometrics, 33(3), 341-365.

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

Nobiling T, Spohn C, Delone M (1998) A tale of two counties: unemployment and sentence severity. Justice Q 15:459–485

Reitz, K. R. (2009). Demographic impact statements, O'Connor's warning, and the mysteries of prison release: Topics from a sentencing reform agenda, 61 Florida Law Review 683.

Rehavi, M. M., and Starr, S. B. (2014). Racial Disparity in Federal Criminal Sentences. Journal of Political Economy, 122(6), 1320-1354.

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

Spohn, C. (2000). Thirty years of sentencing reform: The quest for a racially neutral sentencing process. In Criminal justice 2000: Politics, Processes, and Decisions of the criminal justice system (pp. 427-501). Washington, DC: National Institute of Justice/National Criminal Justice Reference Service.

Spohn, C. (2009). How Do Judges Decide? The Search for Justice and Fairness in Punishment (2nd ed.). Thousand Oaks, CA: Sage.

Spohn, C. and Cederblom, J. (1991). Race and disparities in sentencing: A test of the liberation hypothesis. Justice Quarterly 8:305-27.

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

Starr, S. (2015). Estimating gender disparities in federal criminal cases. Am. L. & Econ. Rev. 17, no. 1 (2015): 127-59.

Steffensmeier, D. and Demuth, S. (2000). Ethnicity and sentencing outcomes in US federal courts: Who is punished more harshly? American Sociological Review, 705-729.

Steffensmeier, D., & Demuth, S. (2001). Ethnicity and judges' sentencing decisions: Hispanic-black- white comparisons. Criminology, 39(1), 145-178.

Steffensmeier, D., Ulmer, J., and 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), 763-797.

Tonry, M. (1995). Malign Neglect: Race, Crime, and Punishment in America. Oxford University Press.

Tonry, M. (1996). Sentencing Matters. Oxford University Press.

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

Ulmer, J. T. (1997). Social worlds of sentencing: Court communities under sentencing guidelines. SUNY Press.

Ulmer, J. T., Eisenstein, J., & Johnson, B. D. (2010). Trial penalties in federal sentencing: extra-guidelines factors and district variation. Justice Quarterly, 27(4), 560-592.

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

Warren, P., Chiricos, T., and Bales, W. (2012). The imprisonment penalty for young Black and Hispanic males a crime-specific analysis. Journal of Research in Crime and Delinquency, 49(1), 56-80.

Williams, R. (2012). "Using the margins command to estimate and interpret adjusted predictions and marginal effects." The Stata Journal 12(2):308-331.

Wooldredge, J. D. (2007). Neighborhood effects on felony sentencing. Journal of Research in Crime and Delinquency, 44(2), 238-263.

Zatz, M. S. (2000). The convergence of race, ethnicity, gender, and class on court decisionmaking: Looking toward the 21st century. In Criminal justice 2000: Politics, Processes, and Decisions of the criminal justice system (pp. 503-552). Washington, DC: National Institute of Justice/National Criminal Justice Reference Service.

Zorn, C. J. (1996). Evaluating zero-inflated and hurdle Poisson specifications. Midwest Political Science Association, 18(20), 1-16.

APPENDIX

Table A1. Comparison of HRM-NB and OLS Models for y > 0							
	HRM-NB Outcome: Prison Term				Log-Linear OLS Outcome: ln(Prison Term)		
	coef.	S.E.	z		coef.	S.E.	z
Case Characteristics							
Offense Seriousness	0.69	0.02	35.91		0.59	0.02	35.55
Commitment Score	0.10	0.01	18.53		0.10	0.00	20.49
Drug Offense	-0.27	0.05	-5.37		-0.16	0.05	-3.15
Property Offense	-0.36	0.03	-13.77		-0.26	0.03	-8.56
Other Offense	-0.40	0.04	-9.44		-0.31	0.03	-9.17
Trial	0.95	0.05	20.17		1.03	0.06	16.83
Mandatory Minimum	0.94	0.04	21.60		1.09	0.05	20.54
Offender Characteristics							
Criminal History	0.01	0.02	0.72		0.02	0.01	2.05
Male	0.08	0.04	2.08		0.07	0.04	1.76
Age	-0.00	0.01	-0.12		-0.01	0.01	-1.25
Age^2	0.00	0.00	1.07		0.00	0.00	2.03
Black	0.39	0.09	4.52		0.32	0.08	3.84
Liberation Hypothesis							
Black x Seriousness	-0.11	0.02	-5.94		-0.09	0.02	-4.75
Black x History	0.00	0.02	0.07		-0.01	0.02	-0.33
Intercept	0.26	0.17	1.52		0.37	0.16	2.30
Adjusted R^2		0.72				0.66	

Notes: Judges' sentencing decisions (prison term in months) modeled using bootstrapped standard errors with 1,000 replications clustered by sentencing judge (51 in total) for both models. With N = 6,611, values of $|z| > 1.96$ are significant at $p < .05$. Estimates are not directly comparable between models because the scale of the dependent variable is different.

STATA CODE

```
* HRM Part 1: Binary (logit) Model Predicting 0s
bootstrap, reps(1000) cluster(judge): logit dv1 c.serious##i.black commitment i.offense i.trial
i.manmin c.history##i.black c.age##c.age i.male
logit dv1 c.serious##i.black commitment i.offense i.trial i.manmin c.history##i.black
c.age##c.age i.male, nolog vce(cluster judge)
* Use Joseph Hilbe's GLME3_software: http://works.bepress.com/joseph_hilbe/60/
abich
margins, dydx(*) noesample vce(unconditional) post
coefplot, horizontal xline(0) yscale(reverse) recast(scatter)  cismooth grid(none)  ///
    order(1.manmin 1.trial serious commitment 4.offense 2.offense 3.offense  ///
    history  1.male 1.black  age )  ///
    coeflabel(1.manmin="Mandatory Minimum" 1.trial="Trial" history="Criminal History"  ///
    serious="Offense Seriousness" 1.male="Male" 1.black="Black"  ///
    commitment="Commitment Score" age="Age" 4.offense="Other Crime"  ///
    2.offense="Drug Crime" 3.offense="Property Crime", wrap(20))  ///
    headings(1.manmin="{bf:Case Characteristics}" history="{bf:Offender Characteristics}")  ///
               xtitle("Average Marginal Effect: Pr(Prison)")

logit dv1 c.serious##i.black commitment i.offense i.trial i.manmin c.history##i.black
c.age##c.age i.male, nolog vce(cluster judge)
margins, dydx(black) at(history = (1 (1) 5)) noesample vce(unconditional) post
coefplot, horizontal xline(0) yscale(reverse) recast(scatter)  cismooth grid(none)  ///
    coeflabel(1._at="None" 2._at="Minimal" 3._at="Moderate" ///
    4._at="Considerable" 5._at="Extensive", wrap(20)) ///
    eqrename(1.black = "Criminal History") eqstrict ///
    xtitle("Discrete Change in Pr(Prison > 0) for Blacks")
fitstat


* HRM Part 2: Zero-Truncated Negative Binomial (ZTBN) Model Predicting Positive Counts
bootstrap, reps(1000) cluster(judge): tnbreg dv1 c.serious##i.black commitment i.offense i.trial
i.manmin c.history##i.black c.age##c.age i.male if dv1>0
tnbreg dv1 c.serious##i.black commitment i.offense i.trial i.manmin c.history##i.black
c.age##c.age i.male if dv1>0, vce(cluster judge)
* Use Joseph Hilbe's GLME3_software: http://works.bepress.com/joseph_hilbe/60/
abich
margins, dydx(*) noesample vce(unconditional) post
coefplot,  horizontal xline(0) yscale(reverse) recast(scatter)  cismooth grid(none)  ///
    order(1.trial 1.manmin serious commitment 2.offense 3.offense 4.offense  ///
    1.male history age 1.black)  ///
    headings(1.trial="{bf:Case Characteristics}" 1.male="{bf:Offender Characteristics}")  ///
    coeflabel(1.trial="Trial" 1.manmin="Mandatory Minimum" ///
    history="Criminal History" serious="Offense Seriousness"  ///
    1.male="Male" 1.black="Black" commitment="Commitment Score" ///
    age="Age" 4.offense="Other Crime"  ///
    2.offense="Drug Crime" 3.offense="Property Crime", wrap(20))  ///
    xtitle("Average Marginal Effect: Predicted Sentence")

tnbreg dv1 c.serious##i.black commitment i.offense i.trial i.manmin c.history##i.black
c.age##c.age i.male if dv1>0, vce(cluster judge)
margins, dydx(black) at(serious = (1 (1) 5)) noesample vce(unconditional) post
coefplot, horizontal xline(0) yscale(reverse) recast(scatter)  cismooth grid(none)  ///
    coeflabel(1._at="Misdemeanor" 2._at="F Felony" ///
    3._at="E Felony" 4._at="D Felony" 5._at="A/B/C Felony", wrap(20)) ///
    eqrename(1.black = "Offense Seriousness") eqstrict ///
    xtitle("Discrete Change in Predicted Prison Term (in Months) for Blacks")
fitstat
```