2016

Examining Prison Effects on Recidivism: A Regression Discontinuity Approach

Ojmarrh Mitchell, Joshua C. Cochran, Daniel P. Mears and William D. Bales

The version of record can be found at https://www.doi.org/10.1080/07418825.2016.1219762.

PRINT VERSION CITATION: Mitchell, Ojmarrh, Joshua C. Cochran, Daniel P. Mears, and William D. Bales. 2017. "Examining Prison Effects on Recidivism: A Regression Discontinuity Approach." Justice Quarterly 34(4):571-596.

PRE-PRINT VERSION

EXAMINING PRISON EFFECTS ON RECIDIVISM: A REGRESSION DISCONTINUITY APPROACH*

Ojmarrh Mitchell
Joshua C. Cochran
Daniel P. Mears
William D. Bales

* Direct correspondence to Ojmarrh Mitchell, PhD, Associate Professor in the University of South Florida's Department of Criminology, 4202 E. Fowler Ave., SOC 107, Tampa, FL 33602, e-mail (omitchell@usf.edu). Joshua C. Cochran, PhD, is an assistant professor at the School of Criminal Justice, University of Cincinnati. Daniel P. Mears, PhD, is the Mark C. Stafford Professor of Criminology at Florida State University's College of Criminology and Criminal Justice. William D. Bales is a Professor at Florida State University's College of Criminology and Criminal Justice.

Examining Prison Effects on Recidivism: A Regression Discontinuity Approach

Abstract

The "get-tough" era of punishment led to exponential growth in the rate of incarceration in the United States. Recent reviews of the literature indicate, however, that limited rigorous research exists examining the effect of imprisonment on the likelihood of future offending. As a result, scholars have called for assessment of this relationship, while using methodologies that can better account for selection effects. This study addresses these calls directly by applying regression discontinuity, a methodology well suited to account for selection bias, on a cohort of felony offenders in Florida. Results suggest that prison, as compared to non-incarcerative sanctions, has no appreciable impact on recidivism. Although no differential effects surfaced across race/ethnicity, the analyses indicated that imprisonment exerts a differential effect by gender with the effect being more criminogenic among males than females.

Keywords: regression discontinuity design; prison; corrections; specific deterrence; recidivism

Examining Prison Effects on Recidivism: A Regression Discontinuity Approach

Limited knowledge exists about the impacts of incarceration on recidivism (Gendreau et al., 2000; Nagin, Cullen, & Jonson, 2009; Villettaz, Killias, & Zoder, 2006). This situation is anomalous considering that, over the past 40 years, U.S. imprisonment rates have increased on a nearly annual basis, yielding rates of imprisonment that make the United States the world's leader in incarceration (National Research Council, 2014; Tonry, 2004). Although the goals associated with imprisonment include retribution, incapacitation of active offenders, and general deterrence of would-be offenders, reduced recidivism constitutes a critical motivation for increasing the use of incarceration in the United States (Cochran, Mears, & Bales, 2014; Loughran et al., 2009; Nagin et al., 2009; Western, 2006).

Accordingly, research that assesses the effectiveness of prison on reoffending is important for understanding the potential benefits, or harms, of large-scale prison expansion. Such scholarship is also important for advancing punishment theory. Competing arguments exist about the impacts of a stay in prison. Rational choice theories, for example, posit that incarceration reduces offending by altering the balance of perceived costs and benefits of criminal behavior (Beccaria, 1963 [1764]; Bentham, 1988 [1789]; see generally, Paternoster, 1987; Apel, 2013). By contrast, learning and labeling perspectives hypothesize that prisons increase offending by serving as "schools of crime" or by stigmatizing offenders in ways that increase their propensity for future criminal behavior (Bernburg & Krohn, 2003; Clemmer, 1940; Irwin, 2005; Paternoster & Iovanni, 1989; Sykes, 1958; Wacquant, 2001).

Scholarship to date has provided limited information about prison's impacts on recidivism. Recent reviews indicate that a stay in prison, as compared to experiencing alternative sanctions like probation, yields null or sometimes harmful, recidivism-increasing effects (Barrick, 2014;

Nagin et al., 2009; Villettaz et al., 2006). Perhaps more troubling, these reviews have found that few methodologically rigorous studies of prison effects on recidivism exist. For example, as many scholars have emphasized (e.g., Cochran et al., 2014; Nagin et al., 2009), most extant studies utilize designs that cannot convincingly minimize the possibility of selection bias. Nagin et al.'s (2009) review found that many of the few stronger study designs are dated and examine sentence lengths that are substantially shorter than those typically found in the United States. In addition, extant work has not systematically assessed whether prison effects vary across two important demographic divides—race and ethnicity, on the one hand, and gender, on the other hand. The disproportionate use of incarceration with minorities and the escalating use of incarceration with females underscore the importance of understanding how, if at all, prison differentially affects some demographic groups more than others.

The goal of this current paper is to address these gaps and limitations in prison research in several ways. First, the study provides a methodologically rigorous assessment of prison effects. Specifically, it employs a research design, regression discontinuity, that has the potential to provide high levels of internal validity (see, e.g., Berk, Ahlman, & Kurtz, 2010; Dunning, 2012; Murnane & Willett 2011; Thistlethwaite & Campbell, 1960). Second, the study assesses the effects of more typical lengths of imprisonment in the United States (at least a year), utilizes a substantially large and comprehensive statewide dataset, and compares prison effects to a realistic counterfactual scenario of non-prison sanctions (i.e., probation, intensive probation, and jail). Third, it assesses whether prison exerts a differential effect for men and women and across racial and ethnic groups. To this end, we begin first by discussing mass incarceration in the United States and then by reviewing empirical research on the effect of imprisonment on recidivism. After identifying findings and limitations in prior work, we describe the analytic approach, regression discontinuity, and the data used in this study. We then discuss the findings and their implications.

BACKGROUND

Mass Incarceration in the United States

Imprisonment in the United States is remarkably common and is especially so for certain demographic groups. The overall U.S. imprisonment rate in 2012 was 471 per 100,000 residents; if those incarcerated in jails are included, the rate increases to 707 per 100,000 residents (Glaze & Herberman, 2013). These rates exceed those of any other nation in the world. For example, the National Research Council (2014) reports that in 2012 the average rate of incarceration (prison and jail confinement) in Western Europe was 100 per 100,000 residents; no other nation had an incarceration rate of more than 500 per 100,000 residents (p. 37). The U.S. incarceration has steadily risen over time. Prior to 1972, for example, the U.S. incarceration rate was never higher than 139 inmates per 100,000 residents and typically fluctuated around 110 inmates per 100,000 residents (National Research Council, 2014). Yet, between 1973 and 2012, the rates of incarceration and imprisonment steadily increased, resulting in the release of approximately 600,000-700,000 prisoners annually.

This situation has led many scholars and policymakers to call for assessments of incarceration to understand better its implications for society (Cullen, Jonson, & Nagin, 2011; Durlauf & Nagin, 2011; Garland, 2001). Of particular salience is the question of whether incarceration in fact reduces recidivism. If it does, that would suggest that, all else equal, imprisonment might be an appropriate policy for public safety purposes. On the other hand, if, as prior work suggests, incarceration may have no effect or in fact may increase recidivism (see, e.g., Cochran et al., 2014; Cullen et al., 2011; Nagin et al., 2009), that would suggest grounds for reconsidering how best to punish offenders while achieving other punishment goals, such as retribution.

Limitations in Prior Research on Prison Effects

The construction of prison facilities in the United States and the expanded use of prisons over the past four decades has been predicated on the idea that incarceration reduces an offender's likelihood of future crime (Cullen et al., 2011). Recent reviews of empirical assessments of prison effects raise critical questions, however, about the ability of incarceration to reduce recidivism successfully compared to other sanctions (Nagin et al., 2009; Smith, Goggin, & Gendreau, 2002; Villettaz et al., 2006). Typically, these reviews find that offenders who were in prison exhibit greater levels of reoffending in comparison to offenders who received non-prison sanctions, but most often these differences are not statistically significant.

For example, Smith et al. (2002) identified 31 studies that collectively produced 104 effect sizes that examined the relationship between imprisonment and subsequent offending. The average phi effect size was 0.07, which indicates a weak, positive relationship with offenders who were imprisoned having marginally greater levels of reoffending. Similarly, a more recent review by Villetaz and colleagues (2006) identified 27 comparisons of prison versus non-prison sentences. Fourteen of these 27 comparisons found no statistically significant difference, 11 found former inmates exhibited higher levels of reoffending, and the two remaining comparisons suggested lower levels of reoffending for former inmates. In short, the existing literature finds that prison generally does not reduce reoffending and in some studies has been found to increase the likelihood of reoffending.

Across these reviews, scholars consistently indicate that the overall number of studies assessing prison effects is relatively small. More importantly, among existing studies, scholars identify at least two critical limitations that are at the same time methodological and theoretical in nature and that weaken claims about prison effects on reoffending.

First, few efforts to date well account for selection bias on unobserved variables. The research designs best suited to deal with selection bias are randomized experiments, natural experiments, and regression discontinuity designs, as each of these designs has the potential to equate imprisoned offenders to non-imprisoned offenders on both observed and unobserved variables (Shadish, Cook, & Campbell, 2001). Yet, these designs have rarely been employed in studies of prison effects. For example, Nagin and colleagues (2009) found only 5 studies that employed rigorous designs; these included 4 randomized experiments and 1 natural experiment. Even these studies are problematic, however. Three of the five studies examined extremely short prison terms (14 days or less). One of the two remaining studies is problematic because it tracked the recidivism of juvenile offenders while the imprisoned group was still confined (Barton & Butts 1990). And one study used a limited sample that focused only on serious, adult felony offenders (Bergman 1976). This situation led Nagin and colleagues (2009) to conclude: "Despite a growing literature on the effect of imprisonment on reoffending . . . rigorous scientific knowledge is in short supply" (p. 116).

Second, despite theoretical reasons to anticipate that prison may exert a differential effect across racial and ethnic groups and among males and females, the existing literature largely has not investigated this possibility (see, however, Mears, Cochran, & Bales, 2012). Instead, studies typically have estimated general effects of imprisonment on recidivism by pooling all cases together, thus implicitly making the assumption that the imprisonment effect is consistent across groups.

As others have suggested, this omission is troubling because some minority offenders are more likely to experience prison and because rates of incarceration of females has been increasing rapidly in recent years (Carson, 2014; National Research Council, 2014). To illustrate, among minority populations, especially in Black communities, imprisonment has become substantially more commonplace—22.4 percent of black men born between 1965 and 1969 have experienced prison whereas only 3.2 percent of white men have had this experience (Western, 2006). In addition, the experience of prison may differ across different demographic groups, which may result in different effects on the probability of reoffending. For example, women may be more adversely affected by the social isolation imposed by prison stays (Holsinger, 2014; Mears et al., 2012). This increasing prevalence and increased contact with formerly incarcerated individuals within minority communities may, among other influences, deteriorate the deterrent effect of prison on minority groups (e.g., Nagin et al., 2009). In addition, prison experiences may more adversely affect minority inmates if they experience greater levels of discrimination and social isolation or have more diminished access to programming (e.g., Hemmens & Stohr 2014).

REGRESSION DISCONTINUITY DESIGN

The goal of this study is to address these two limitations. We attempt to address the first limitation by applying a research methodology—regression discontinuity (RD) design—that is commonly used in many areas of policy evaluation and has been recommended in prior reviews, yet is rarely used in punishment scholarship. As we discuss below, the RD design has the potential to account better for unobserved characteristics that might lead to selection bias. To address the second limitation, we first undertake an assessment of prison effects on recidivism using a general sample of imprisoned felony offenders, and then we repeat the assessment after splitting the sample into separate race and ethnic groups, and, separately, into a sample of males and a sample of females. We began by describing the utility of the regression discontinuity design for contemporary punishment research, along with the analytic strategy that we employ.

Randomized experiments are referred to as the scientific "gold standard" because of their high internal validity. The reason is that assignment to the treatment group is exogenously determined by randomization. Respondents in the treatment and control groups, respectively, are equal in expectation on all variables, even those that are unobserved (Dunning, 2012; Murnane & Willet 2011; Shadish et al., 2001). Randomized experiments, however, are impractical or unethical in many situations, particularly those related to serious offenders and their punishment.

In lieu of experimental designs, scholars have emphasized the strengths of RD as an alternative means of obtaining unbiased estimates of treatment effects (Thistlethwaite & Campbell, 1960). In the past decade, the RD design has become increasingly recognized by and used in economic analyses of policies and programs (see Cook, 2008). To our knowledge, however, no studies have employed the design to utilize sentence scores and estimate the effects of prison. However, studies by Loeffler and Grunewald (2015) and Lee and McCrary (2009) have successfully used age boundaries to estimate the effect of processing "juveniles" in the adult criminal justice system; and, Chen and Sharpiro (2007) used the RD design to estimate the effect of prison conditions (i.e., higher security levels) on recidivism. Indeed, more than 50 years after its discovery, RD has been infrequently applied to crime and justice policies. In fact, in criminology and criminal justice journals, only a small number of studies have been published using RD (see, e.g., Berk et al., 2010; Jalbert et al., 2010; Loeffler & Grunwald, 2015; Maddan et al., 2011; Rauma & Berk, 1987; Worrall & Morris, 2011).

Whatever the reason, the situation is unfortunate, given the many similarities between RD and randomized experiments. Perhaps the most important similarity between the two is their high level of internal validity (Berk et al., 2010; Dunning, 2012; Murnane & Willett, 2011). Randomized experiments and RD have high levels of internal validity because both methods assign respondents to conditions exogenously. Essentially, randomized experiments use a random variable to assign respondents to conditions. Somewhat similarly, RD uses a non-random "rating variable" and a cut score to assign respondents to conditions. Respondents scoring above the cut score on the rating variable are assigned to one condition, typically the treatment, and those below the cut score are assigned to another group, typically the control. Therefore, assignment in RD is exogenous, *conditional on the rating variable*, which can result in a situation in which respondents near the cutoff score are comparable on observed and unobserved variables (Dunning, 2012; Murnane & Willet, 2011).

Another similarity between randomized experiments and RD is that assignment is often imperfect. In nearly all randomized experiments, some respondents assigned to the treatment condition do not actually receive the prescribed treatment and some respondents assigned to the alternative condition may end up receiving the treatment. Likewise, in RD, some respondents with scores above the cut-point may not actually receive the treatment and some respondents below the cut-point will receive treatment. When this situation occurs, the discontinuity is referred to as a "fuzzy" discontinuity, as opposed to a "sharp" discontinuity.

In randomized experimental designs, this situation can be accounted for by conducting intent-to-treat analyses to preserve the exogeneity established by random assignment. Similarly, as we illustrate in the analyses below, the exogeneity created by a rating variable and its cut score can be preserved through use of the rating variable's assignment rule—regardless of whether the actual treatment received matched the assignment rule. These kinds of analyses preserve the exogeneity created by the assignment procedure; however, these analyses underestimate the true treatment effect because many of those assigned to receive the treatment do not actually receive the treatment. In both randomized experiments and RD, the effect of the treatment received, as opposed to treatment as assigned, can be estimated by using the intended treatment status variable as an instrumental variable and two-stage regression (Dunning, 2012; Murnane & Willett, 2011).

DATA

The data for this study come from the Florida Department of Corrections (FDOC) database. These data consist of individuals who were convicted of felonies between 1999 and 2002 in any of Florida's 20 judicial circuits and were sentenced to probation, intensive probation, jail, or prison (N = 330,971). This data set combines information regarding sentencing (e.g., total sentence points) with information in the FDOC's Offender Based Information System (OBIS), which records information regarding prior record (e.g., prior prison commitments, supervision violations) and information on prison release dates. Recidivism was linked to each offender, and is defined as a felony reconviction within a three-year follow-up period. For prison and jail, the recidivism "clock" starts upon release from incarceration; for either type of probation, it starts when the sentence begins.[1]

ANALYTIC STRATEGY

The analyses below aim to address methodological limitations and to extend research on punishment effects in several ways. First, we utilize a research methodology, RD, well-equipped to address selection bias. Second, we use a contemporary, statewide sample of adult felons sentenced to prison for at least a year and a day and of felons who received a non-prison

---

[1] Because the recidivism clock starts immediately after sentencing for those sentenced to a non-prison sanction but starts at least a year after sentencing for those sentenced to a prison sanction, it is possible that those sentenced to prison will be systematically older than those sentenced to probation. However, as we demonstrate below, there is no substantively or statistically significant difference between the two groups' age at the beginning of the recidivism-tracking period.

sanction. This approach addresses limitations in some prior studies, which typically have utilized small, dated, and non-representative samples. Third, we examine variation in the effect of a prison sentence on recidivism by gender and race/ethnicity.

Total Sentence Points as a Rating Variable

We use Florida's Criminal Punishment Code (CPC) along with its "total sentence points" as a rating variable. Beginning in October of 1998, Florida established CPC as a guide to sentencing all felony offenses with the exception of capital felonies (Florida Criminal Punishment Code, 2012). Under the CPC, elements of the offense, including seriousness of the offense, victim injury, defendant's prior criminal record, defendant's legal status at the time of the offense (e.g., community supervision), use of firearm, and so forth, are scored to compute total sentence points. Total sentence points establish "the lowest permissible sentence" under the CPC. "The lowest permissible sentence is the minimum sentence that may be imposed by the trial court, absent a valid reason for departure" (Florida Criminal Punishment, Code 2012, p. 13). A prison sentence is the lowest permissible sentence for cases with total sentence points greater than 44 points. The lowest permissible sentence in cases scoring 44 points or less is a non-prison sanction, which may include probation, jail, community control (i.e., intensive probation with house arrest or curfew), or a combination of these non-prison sanctions. Thus, in Florida, a case with more than 44 total sentence points is said to have "scored to prison"; we use this term in the remainder of this research to refer to cases intended to receive a prison sentence based on the CPC.

Total sentence points acts as a rating variable and the CPC's rules regarding the lowest permissible sentence establish a cut score useful for estimating the effect of prison in comparison to non-prison sanctions. For our purposes, we consider cases scored to prison (i.e., more than 44 total sentence points) as being assigned to the "treatment" (a prison sentence) and not scored to prison (i.e., 44 total sentence points or less points) as being assigned to "no treatment" (a non-prison sanction). The CPC also guides the length of prison sentences in cases scored to prison. Specifically, according to the CPC, prison sentences are determined by subtracting 28 points from the total sentence points and then decreasing the remaining total by 25 percent, and all prison sentences must exceed one year. Thus, cases just above the cut-point of 44 points are intended to receive a prison sentence of just more than a year (e.g., a year and a day). However, the CPC allows judges to depart from the lowest permissible sentence in certain situations, which, as we demonstrate below, makes the discontinuity "fuzzy."

Valid application of the RD design requires that the rating variable cause a discontinuity in receipt of the treatment of interest at the rating variable's cut-point. Below, and as a preliminary step in the analysis, we demonstrate that there is clear evidence of a large discontinuity in the assignment of cases to prison at the cut score established by Florida's CPC. Valid application of the RD design also requires that cases "near" the cut score are highly comparable on observed variables. Bloom (2012) notes that this is the "most important test of the internal validity of a regression discontinuity estimation model. … If a regression discontinuity estimation model is internally valid, there should be few, if any large or statistically significant treatment/control group baseline discontinuities" (p. 63). Using bivariate and regression (to condition on the rating variable) analyses, we demonstrate that cases near the cut score in fact are comparable.

Modern Regression Discontinuity Design

Modern regression discontinuity analysis (see e.g., Bloom, 2012; Dunning, 2012; Murname

&Willett, 2011) conceptualizes the RD design as equating cases near the cut score.[2] This conceptualization of the RD design leverages "local randomization" among cases that reside immediately on each side of a given cut score. It follows from this conceptualization that analyses are restricted to cases neighboring the cut score—i.e., "local" analyses. The local randomization that occurs around the cut score allows us to use nonparametric analyses similar to those used in the analysis of randomized experiments, yet in RD analysis we typically must control for the scores on the rating variable. Generically, this approach can be represented by the following regression model:

$$Y_i = \alpha + \beta_0 T_i + \beta_1 r_i + \beta_2 r_i * T_i + \varepsilon_i$$

Here, Y_i is the outcome variable of interest for observation i, T_i, is the dichotomous variable flagging those assigned to the treatment, r_i is the rating variable which has been centered at the cutoff score and may include interactions with T_i, ε_i is the random error, and the β_0 is the effect of being assigned to the treatment. Given that the outcome variable of interest, three-year reconviction, is dichotomous, we use probit regression with robust standard errors within the specified bandwidths (i.e., local analyses).[3] The analysis described above was conducted on all cases and then disaggregated by race/ethnicity, and gender to investigate variation in the effect of scoring along these factors.

The analytic strategy discussed above accurately estimates the effect of *scoring to prison* on reconviction; however, this method certainly underestimates the effect of *actual imprisonment* on

[2] By contrast, traditional applications of RD design focus on examining a discontinuity in an estimated regression line depicting the parametric relationship between the outcome and the rating variables, hence the name "regression discontinuity."
[3] We use robust standard errors to correct for clustering of observations within judicial circuits. To assess the appropriateness of our application of robust standard errors, we conducted sensitivity analyses following King and Roberts's (2015) approach. Specifically, we compared the robust standard errors from the probit models to standard errors for random-effects probit models, which explicitly account for clustering. We found very small differences between the two kinds of standard errors (the largest difference was .001); the small magnitude of these differences indicates that our use of robust standard errors is appropriate.

reconviction due to the fact that a large percentage of cases scoring to prison did not actually receive a prison sentence. To examine the effect of actually being imprisoned, as opposed to scoring to prison, we use instrumental variable estimation and two-stage regression—a technique that has become standard in RD analyses with fuzzy discontinuities (see, e.g., Dunning, 2012; Imbens & Lemieux, 2008; Murnane & Willet, 2011). Generally, in this approach the endogenous variable (actual imprisonment, in this case) is regressed on the instrumental variables (scoring to prison, in this case) to identify exogenous variation in the endogenous variable. The predicted values of this model are used in the second stage model, where the dependent variable of interest (reconviction) is regressed on these predicted values to estimate the causal effect of the endogenous variable on the dependent variable. Both of these stages are estimated simultaneously using two-stage regression. However, this approach is known to be inconsistent, if the endogenous variable is dichotomous like our actual imprisonment decisions variable. Applying instrumental variable analysis with a dichotomous endogenous variable is known as a "forbidden regression" model. Wooldridge (2002, p. 623-625) offers a solution to forbidden regression model problem, which involves three steps: (1) the endogenous variable is regressed on the instrumental variables using a probit model to obtain the predicted values of this model; (2) the endogenous variable is regressed on the predicted values and exogenous variables using a OLS regression to obtain the predicted values of this model; and (3) the dependent variable of interest, reconviction, is regressed on the predicted variables from step 2 along with the exogenous variables. We applied this process, which has been successfully applied elsewhere (Adams, Almeida, & Ferreira, 2009). The instrumental variable approach provides a more accurate estimate of the effect of an actual prison sentence; however, because this technique utilizes only the exogenous part of actual prison sentences (i.e., the part that can be predicted by scoring to prison) there is a loss of statistical power. Here, again, this analysis was conducted on all cases and then disaggregated by race, ethnicity, and gender.

Selecting Bandwidth Size Around the Cut Score

Before moving forward with the analyses, we have to determine the proper bandwidth size around the cut score to be used to create the localized sample. This is a critical step in the modern RD design and a central issue in local linear regression analysis (see, e.g., Imbens & Kalyanaraman, 2009). Narrower bandwidths maximize the comparability between those on either side of the cut score but it simultaneously reduces statistical power. On the other hand, wider bandwidths maximize the precision of estimates (i.e., produce smaller standard errors and greater statistical power) but raise concerns about internal validity due to the inclusion of cases farther from the cut score. We utilized an algorithm developed by Imbens and Kalyanaraman's (2009) to select the bandwidth minimizing bias and optimizing precision. The optimal bandwidth was 1.7 points; that is, cases within 1.7 points of the cut score in either direction were defined as "near" the cut score and included in the local linear regression analyses. Thus, the local linear regression analyses were restricted to cases with total sentence points between 42.3 and 45.1. Notably, there is a statistically powerful number of cases in this bandwidth (n = 11,047). To test the robustness of our findings to different bandwidths, we report all results using the optimal bandwidth (100% of the optimal bandwidth), half the optimal bandwidth (50% of the optimal bandwidth), and twice the optimal bandwidth (200% of the optimal bandwidth).

Altering the size of the bandwidth not only serves as a sensitivity test for the estimated effects, it also serves as a check against finding statistical significance by chance. Given the relatively large number of comparisons, the probability of making a Type I error is non-trivial. As a guard, we interpret with caution effects that are statistically significant in only one of the three bandwidths. Last, there is a longstanding debate in the methodological literature concerning the drawbacks of null hypothesis statistical testing, particularly when multiple comparisons are being made (for recent commentary, see Gelman and Loken, 2013, 2014). Rather than focusing solely on statistical significance, we also estimate the marginal change in the probability of reconviction for those assigned to prison in comparison to those not assigned to prison. These estimates allow readers to look beyond statistical significance to consider the substantive significance of the findings and to judge whether the differences are practically meaningful.

RESULTS

Preliminary RD Analyses

For the 330,971 cases contained in these data, total sentence points ranged from .2 to 5,148 points with a heavy right skew. The mean total sentence points was 33.5 (*SD* = 30.1) and the median total sentence points was 28. Thus, most cases had total sentence points below the cut-point of 44 points. In fact, 44 total sentence points roughly corresponds to the 80th percentile.

Figure 1 demonstrates that, as expected, there is a fuzzy discontinuity in the probability of receiving a prison sentence at the cut-point specified by Florida's CPC. For illustrative purposes, Figure 1 examines the probability of imprisonment for cases with total sentence scores within two optimal bandwidths of the cut score, that is, 40.6 to 47.4 points. We choose to focus on this relatively wide range of sentence scores because here we seek only to identify whether there is a discontinuity near the cutoff score, not to estimate the effect of the assignment rule.

This figure shows that there is an obvious and sizeable increase in the probability of imprisonment for cases above the cut score in comparison to the cases below the cut score. We conducted a probit regression analysis (not shown) that regressed the imposition of a prison sentence on the dichotomous variable flagging cases scored to prison within two optimal bandwidths of the cut score. This model indicated that the propensity for having a prison sentence imposed instead of a community based sanction is .944 ($z = 47.52$, $p < .001$) greater for those scored to prison (greater than 44 sentencing points) compared to those not scored to prison (44 points or less). When the focus is narrowed to cases within the optimal bandwidth of the cut score, the propensity for having a prison sentence imposed is .850 ($z = 30.50$, $p < .001$) greater for those scored to prison. These results indicate that, although it may be "fuzzy," there is a clear discontinuity in the imposition of a prison sentence at the cut score specified by Florida's CPC.

Insert Figure 1 about here.

One of the threats to RD design is manipulation of scores on the rating variable. Various actors may manipulate rating variable scores in one direction or another to move cases below or above the cut score in an effort to evade the assignment rule. McCrary (2008) provides a commonly utilized test for manipulation of continuous rating variables that compares the density of scores on both sides of the cut score. According to the logic of this test, if the density of the rating variable changes at the cut score, then this is evidence of manipulation. Frandsen (2014) provides a modification to McCrary's test for discrete rating variables like total sentence points, which is measured in .1 increments. Figure 2 displays a kernel density plot of total sentence points. Visual inspection of this plot clearly suggests that total sentence points were manipulated downward (below the 44 cut score) in an effort to avoid scoring to prison. Frandsen's test comports with this visual assessment by finding that the density of the plot differs significantly above and below the cut score ($p < .001$). This finding is not surprising as attorneys, particularly defense attorneys, will work to try to obtain total sentence points that fall below the cut score.

Insert Figure 2 about here.

Manipulation of the rating variable is most damaging to the internal validity of RD designs when it causes those above the cut score to differ from those below the cut score. In fact, Bloom notes that the "most important test of the internal validity of a regression discontinuity estimation model" concerns the comparability of cases just above and below the cut score; "[I]f a regression discontinuity estimation model is internally valid, there should be few, if any large or statistically significant treatment/control group baseline discontinuities" (p. 63).

Simply put, according to the logic of RD, the discontinuity in the imposition of a prison sentence at the cut score is supposed to equate cases scored to prison and not scored to prison near the cut score, respectively, particularly after controlling for the small difference between groups on the rating variable. Here, we define cases near the cut score as those cases within the optimal bandwidth (1.7 points in either direction of the cut score). Table 1 summarizes key demographic, prior record, and offense features to illustrate the usefulness of the rating variable cutoff; notably, *these variables represent every individual-level variable in the data set*. The left side of Table 1 summarizes these measures when the full sample is bifurcated at the sentence score cutoff. Here, we can see that there are sizeable differences between cases scored to prison in comparison to cases not scored to prison. Not surprisingly, these two groups differ statistically ($p < .01$) on all of the variables listed in Table 1.

Insert Table 1 about here.

By contrast, the right side of Table 1, which includes only those cases within the optimal bandwidth, displays much smaller differences between the two groups. Inspection of the variable means reveals that the cases just below and just above the cut score are highly similar on the observed covariates. Specifically, these two groups are nearly identical with respect to age, gender distribution, racial and ethnic distribution, offense type distribution, prior convictions, and prior supervision violations. As expected, the characteristics of the offenders near the cut score indicate that these individuals are generally composed of lower level, nonviolent offenders—more than half of all of these individuals were convicted of property or drug crimes—who have a limited number of prior convictions and prior stays in prison. As we describe in the conclusion section, these are precisely the types of offenders on which current debates about future sentencing policies typically are centered (e.g., Savage, 2013; Apuzzo, 2014).

Although the two groups were largely similar across these characteristics, the two groups exhibited small, but statistically significant differences on two variables: prior prison commitments and prior supervision violations. This finding is not surprising, given the statistical power of these comparisons (n = 11,047) and the fact that by definition those just above the cut score have slightly more criminal history (criminal history is part of total sentence points). Importantly, however, both of these differences were no longer statistically significant once total sentence points was controlled using OLS regression (not shown here to conserve space but available upon request). Specifically, we regressed the number of prior prison commitments on intended assignment (prison or non-prison sentence) and total sentence points. After controlling for total sentence points, the intended sentence variable (prison or non-prison sentence) was not significantly related to number of prior prison commitments. The same analysis was conducted with prior supervision violations as the dependent variable; the result was the same. Thus, the cut score specified by the CPC appears to be effective in equating cases adjacent to the cut-point *across every individual-level variable in the data set*, which supports the validity of this application of the RD design.

Does Scoring to Prison Affect Recidivism?

As discussed above, we utilized local probit regression models to estimate the effect of being scored to prison versus not being scored to prison, while controlling for total sentence score. A scatterplot of the probability of three-year reconviction by total sentence points within two optimal bandwidths of the cut score is displayed in Figure 3. A visual examination of this scatterplot reveals no discernible discontinuity in the probability of reconviction at the cut score; this implies that the effect of scoring to prison on reconviction is not large. Figure 3 also clearly indicates that the roughly 43% of those in this bandwidth were reconvicted within 3 years.

Insert Figure 3 about there.

Table 2 summarizes the results of the probit regression analyses. The reported models regress reconviction on the scored to prison indicator variable in the specified bandwidth, while controlling for total sentence points. These models do not include interactions between scored to prison and total sentence points as this interaction was not statistically significant in any of the models; thus, the relationship between total sentence points and reconviction does not differ for those just above and just below the cut score. These coefficients are reported for all cases in the bandwidth, then disaggregated by race and ethnicity, and again disaggregated by gender. Further, we assess the robustness of these results across bandwidth sizes. These results indicate that in the optimal bandwidth scoring to prison increases the propensity for reconviction by .053 or more intuitively increases the marginal probability of reconviction by roughly 2 percentage points (from 43% to 45%). The evidence here, then, is for a modest difference in the likelihood of reconviction. These models also indicate that as the bandwidth increases the magnitude of scoring to prison's effect shrinks towards zero—this finding of diminishing effects holds for nearly all of the results from the local probit regression models.

The middle portion of Table 2 reports the effect of scoring to prison for Blacks, Hispanics, and Whites. The effect of scoring to prison was strongest among Blacks and smallest among

Hispanics, yet differences by race/ethnicity were not statistically significant. Among Blacks, the effect of scoring to prison was consistently positive indicating increases in the likelihood of reconviction. For example, in the optimal bandwidth the effect of scoring to prison was to increase the propensity of reconviction by .111 units (a 4.4 percentage point increase in the marginal probability of reconviction). Yet, for Hispanics, the effect of scoring to prison was consistently negative but non-statistically significant; for instance, in the optimal bandwidth the predicted difference in the probability of reconviction for those assigned to prison was 9 percentage points lower than those not assigned to prison. Minority and white differences in the effect of scoring to prison were not statistically significant, which indicates again that the effect of scoring to prison does not vary by race or ethnicity. Together, these results suggest that the effect of scoring to prison is invariant among Whites, Blacks, and Hispanics.

Insert Table 2 about there.

We do identify differential effects by gender. Inspection across the two narrower bandwidths reveals that the effect of scoring to prison increases the likelihood of reconviction for males but not for females. In the optimal bandwidth, the effect of scoring to prison was to increase the propensity of reconviction by .098 for males but decrease the propensity of reconviction by .244 for females. Substantively, the predicted probability of males who were assigned to prison are 3.9 percentage points higher than males not assigned to prison; for females, there is a -9.0 percentage point difference indicating that females assigned to prison had lower probabilities of reconviction. Both of these effects were marginally statistically significant ($p < .10$) and differed statistically from one another. Yet, in the widest bandwidth, neither gender-specific effects were statistically significant and the difference in these effects also was not statistically significant. Thus, in the two narrower bandwidths the effect of scoring to prison differed by gender with scoring to prison increasing reconviction for males and decreasing reconviction for females.

<u>Effect of Imprisonment Using Instrumental Variable Estimation</u>

The above analyses use the dichotomous scoring to prison variable and total sentence points to estimate the effect of assignment, or scoring to prison on reconviction. Yet, in many cases, offenders assigned to a prison sentence based on the CPC's policies were sentenced by the court to a non-prison sentence, and some cases assigned to non-prison sanctions based on the CPC's policies were sentenced by the court to a prison sentence. In short, judges used their discretion to depart from the CPC's general assignment rules. The question becomes: What is the effect on recidivism of actually having gone to prison?

To answer this question, we used the instrumental variable models for a dichotomous endogenous variable as described above (see the "Analytic Strategy" section). The results of the final stage model, which estimates the relationship between imprisonment and reconviction are summarized in Table 3 first for all cases then broken down by race and gender. Across these results, we see substantively similar results to those described above in that the direction of the estimated relationships and the pattern of statistical significance are the same in both sets of analyses. However, the estimated relationships are consistently about four times larger than those estimated in the earlier analyses (see Table 2). For example, the earlier results estimating the effect of scoring to prison on reconviction found that the general effect in the optimal bandwidth was .053 but the effect of actual prison placement on reconviction in the optimal bandwidth was .212; yet, because the standard error also grew in magnitude, this effect is not statistically significant. Further, just as in the earlier analyses, the magnitude of this effect diminishes as the bandwidth increases.

The findings from the analyses disaggregated by race/ethnicity also parallel the earlier findings. In particular, these findings indicate that in the two narrower bandwidths, imprisonment increased reconviction for Blacks and this effect was statistical significant at the .10 level; likewise, imprisonment increased reconviction among Whites but this effect was not statistically significant. The difference in the magnitude of these race specific effects was not statistically significant. Notably, the effect of imprisonment on reconviction could not be accurately estimated among Hispanics in the two narrower bandwidths, as the instrumental variable (scored to prison or not) did not predict the imprisonment decisions for Hispanics in these bandwidths. Therefore, the results of the Hispanic specific models in the two narrower bandwidths do not meet the requirements of instrumental variable analysis and are omitted for this reason. In the widest bandwidth (200% of the optimal bandwidth), none of the race/ethnicity specific effects are statistically significant and none of these effects differ statistically for the others.

The gender specific effect analyses continue to indicate gender differences in the effect of sanctioning on the likelihood of reconviction. In the two narrower bandwidths, the effect of imprisonment was to increase the likelihood of reconviction among males but to decrease this likelihood among females, and these effects differ statistically. Only in the widest bandwidth, where the effects of imprisonment on reconviction shrink towards zero, are these effects no longer statistically different.

DISCUSSION AND CONCLUSIONS

Recent reviews and meta-analyses indicate substantial uncertainty about the impact of prison on those who experience it (e.g., Nagin et al., 2009; Villetaz et al., 2002). These assessments indicate that a limited number of methodologically rigorous studies exist and those that do exist are limited by examining short terms of imprisonment or examining the effects of imprisonment prior to the current era of mass incarceration. The lack of rigorous assessment raises concerns given the fiscal and social costs of incarceration and the increase in imprisonment over the past 40 years. In the absence of clear evidence about prison effectiveness and in the face of concerns about these costs, policymakers and practitioners have called for a reevaluation of U.S. incarceration practices (Savage, 2013; Apuzzo, 2014).

The goal of this study was to extend research on prison effects and to better address the limitations identified in prior studies. This research used a methodology—regression discontinuity (RD) design—that accounts for unobserved characteristics that might lead to selection bias better than most existing studies. This analysis also extends research by assessing the effect of a prison term of common length in the United States using a contemporary sample of convicted offenders. Not least, it extends prior work by examining whether prison effects on recidivism vary by gender and by race and ethnicity.

The methodological rigor offered by RD design makes this study an important contribution to the extant research. There is, however, an important caveat that warrants emphasis. The application of RD here utilizes only cases adjacent to the cut score and therefore this study only estimates the "local" treatment effect of imprisonment. In our view, this focus on cases near (local to) the cut score is appropriate because it best simulates "real world" considerations—that is, it is these cases that are near the cut score for which the appropriate type of sentence (e.g., probation versus prison) is most ambiguous. Thus, a focus on such cases provides theoretically- and policy-relevant information in regard to the types of sentences most likely to minimize subsequent offending. Even so, it is important to bear in mind that the estimated effect of a prison sentence on recidivism may be markedly different for cases with sentencing scores substantially higher or lower than the cut score.

Five critical findings emerged from this study. First, the RD design was able to address confounding due to selection bias. To be sure, our application of RD design was not perfect—for example, we found evidence of manipulation of scores on the rating variable. However, this manipulation did not appear to affect negatively the comparability of those around the cut score. Cases on both sides of the cut score were highly similar on measured covariates; in fact, there were no statistically significant differences between the groups on any of the individual-level covariates in the data set after conditioning on the rating variable. The similarity on observed variables suggests that the rating variable also balanced the groups on salient unobserved variables. Thus, one methodological implication of this research is that the RD approach can be applied to answer important criminological questions. Given the widespread use of structured sentencing and risk assessments with cutoff scores, there exist many opportunities to utilize RD designs in assessing prison effects.

Second, findings of the RD analyses support prior theoretical and empirical assessments of prison effects, indicating that imprisonment generally yields no effects or substantively small adverse effects on the likelihood of reconviction compared to alternative sanctions (e.g., Cochran et al., 2014).[4] This finding has straightforward policy implications. Specifically, it raises questions about the utility of imprisonment for offenders of marginal seriousness (i.e., offenders near the prison cut score). The majority of these offenders were non-violent property and drug offenders, who constitute the same offender population that has been at the center of recent policy debates and questions surrounding the future use of imprisonment in the United States. The results here underscore the argument that alternatives to imprisonment may be more useful for these types of offenders.

Third, a corollary of the above findings is that the analyses identified no beneficial effects of

[4] Our finding that imprisonment generally has no deterrent effect and may cause a small increase in recidivism also parallels Chen and Sharpiro's (2007) regarding the effect of harsher prison conditions on recidivism.

incarceration on recidivism. This result raises questions about get-tough policy shifts in the past 40 years that have expanded the range of sanctions that can result in a prison sentence and that have greatly increased the number of individuals admitted to the prison system. There remains little empirical evidence that prison reduces recidivism. It is, of course, possible that incarceration is effective in other ways, such as achieving retribution or greater public safety through other mechanisms, such as general deterrence (Durlauf & Nagin, 2011).

Fourth, we found no variation in the effect of incarceration on recidivism by race/ethnicity. There are, however, reasons to anticipate racial variation. For example, experiences during incarceration may be disparate, such that minorities experience harsher or unfair treatment while incarcerated, which may exert criminogenic effects. Despite the lack of variation in prison effects across race/ethnicity identified here, future research should more systematically test these potential mechanisms and variations in sanction effects across race and ethnicity.

Fifth, our findings reveal that the effect of imprisonment may vary by gender. For males, the imprisonment effect was harmful; prison, as compared to non-prison sanctions, substantially increased the likelihood of recidivism. For females, the imprisonment effect was negative—indicating a reduction in reconviction—but this effect was unreliable and not statistically significant. Although we cannot reject the null hypothesis that imprisonment had no effect on females' likelihood of reconviction, we can conclude that the effect of imprisonment for males appears to increase reconviction and is different for males and females near the prison cut score.

Why does the effect of imprisonment vary by gender? Although we cannot answer this question with these data, there are several reasons to suspect that imprisonment may entail different experiences for females as compared to males. Historically, prisons in the United States were developed with male offenders in mind—the needs of female inmates were largely an afterthought (Holsinger, 2014). Contemporary research suggests that certain aspects of the prison experience differ for women. The prevalence of depression and victimization, for example, is greater among female inmates (Holsinger, 2014). Such differences in turn may accumulate and generate different effects of incarceration on recidivism (Mears et al., 2014).

The implication of these findings, particularly when paired with recent assessments of sanction effects (e.g., Bales & Piquero, 2012; Cochran et al., 2013; Nagin et al., 2009; Nagin & Snodgrass, 2013), are straightforward: The effect of prison on reoffending is unclear and may be null or contribute to more rather than less recidivism. It of course is possible that competing effects exist that offset one another. For example, deterrent effects might well exist but be offset by labeling effects and collateral consequences of punishment (e.g., Bernburg & Krohn, 2003; Hagan & Dinovitzer, 1999). At the same time, the other part of the equation—probation and community sanctions—may influence estimated effects. Non-incarcerative sanctions, for example, may consistently provide offenders the supervision, rehabilitative services, opportunities for employment, and other benefits that may be more likely to deter offenders or, through these other mechanisms, promote desistance.

After nearly four decades of sustained growth in the prison system, evidence suggests that recidivism rates are high (DuRose, Cooper, & Snyder 2014) and compelling theoretical arguments suggest that the experience of incarceration may be criminogenic (Cullen et al., 2011). There remains, however, a remarkable shortage of credible, empirical studies assessing prison effects and their causes. The time is none too soon to correct this situation and determine if, in fact, prison can reduce recidivism and, if it can, the specific conditions under which such a benefit might arise.

REFERENCES

Adams, R., Almeida, H., & Ferreira, D. (2009). Understanding the relationship between founder-CEOs and firm performance. *Journal of Empirical Finance, 16*, 136-150.

Apel, R. (2013). Sanctions, perceptions and crime: Implications for criminal deterrence. *Journal of Quantitative Criminology, 29*, 67-101.

Apuzzo, M. (March 14, 2014). Holder backs proposal to reduce drug sentences. *The New York Times,* p. A18.

Bales, W. D., & Piquero, A. R. (2012). Assessing the impact of imprisonment on recidivism. *Journal of Experimental Criminology, 8*, 71-101.

Barrick, K. (2014). A review of prior tests of labeling theory. In D. P. Farrington & J. Murray (Eds.), *Labeling theory: Empirical tests* (pp. 89-112). New Brunswick, NJ: Transaction Publishers.

Barton, W. H., & Butts, J. A. (1990). Viable Options: Intensive Supervision Programs for Juvenile Delinquents. *Crime & Delinquency, 36*, 238-256.

Beccaria, C. (1963). *On crimes and punishment*. Englewood Cliffs, NJ: Prentice Hall.

Bentham, J. (1988 [1789]). *The Principles of Morals and Legislation*. New York: Prometheus Books.

Bergman, G. R. (1976). *The Evaluation of an Experimental Program Designed to Reduce Recidivism among Second Felony Criminal Offenders*. Ph.D. dissertation, Wayne State University: Department of Measurement and Evaluation.

Berk, R., Barnes, G., Ahlman, L., & Kurtz, E. (2010). When second best is good enough: a comparison between a true experiment and a regression discontinuity quasi-experiment. *Journal of Experimental Criminology, 6*, 191-208.

Bernburg, J. G., & Krohn, M. D. (2003). Labeling, life chances, and adult crime: The direct and indirect effects of official intervention in adolescence on crime in early adulthood. *Criminology, 41*, 1287-1318.

Bloom, H. S. (2012). Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness, 5*, 43-82.

Carson, E. A. (2014). *Prisoners in 2013*. Washington, D.C.: Bureau of Justice Statistics.

Chen, M. K., & Shapiro, J. M. (2007). Do Harsher Prison Conditions Reduce Recidivism? A Discontinuity-based Approach. *American Law & Economics Review, 9*, 1-29.

Clemmer, D. (1958). *The prison community*. New York: Holt, Rinehart, and Winston.

Cochran, J. C., Mears, D. P., & Bales, W. D. (2014). Assessing the effectiveness of correctional sanctions. *Journal of Quantitative Criminology, 30*, 314-347.

Cook, T. D. (2008). "Waiting for Life to Arrive": A history of the regression-discontinuity design in Psychology, Statistics and Economics. *Journal of Econometrics, 142*, 636-654.

Cullen, F. T., Jonson, C. L., & Nagin, D. S. (2011). Prisons do not reduce recidivism: The high cost of ignoring science. *The Prison Journal, 9*(48S-65S).

Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. New York: Cambridge.

Durlauf, S. N., & Nagin, D. S. (2011). Imprisonment and crime: Can both be reduced? *Criminology and Public Policy, 10*, 13-54.

Durose, M. R., Cooper, A. D., & Snyder, H. N. (2014). *Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010*. Washington, D.C.: Bureau of Justice Statistics.

Florida Department of Corrections & Office of the State Courts Administrator. (2012). *Florida Criminal Punishment Code*. , Florida Department of Corrections & Office of the State Courts Administrator. Tallahassee, FL.

Frandsen, B. R. (2014). *Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete*. Manuscript, Brigham Young University, Department of Economics.

Garland, D. (2001). *Mass imprisonment: Social causes and consequences*. Thousand Oaks, CA: Sage Publications.

Gelman, A., & Loken, E. (2013). *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time*. Unpublished manuscript, http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf.

Gelman, A., and Loken, E. (2014). The statistical crisis in science. *American Scientist*. 102:460.

Gendreau, P., Goggin, C., Cullen, F. T., & Andrews, D. A. (2000). The effects of community sanctions and incarceration on recidivism. *Forum of Corrections Research, 12*, 10-13.

Glaze, L. E., & Herberman, E. J. (2013). *Correctional Populations in the United States, 2012*. Washington, D.C.: Bureau of Justice Statistics.

Hagan, J., & Dinovitzer, R. (1999). Collateral consequences of imprisonment for children, communities, and prisoners. *Crime and Justice, 26*, 121-162.

Hemmens, C., & Stohr, M. K. (2014). The racially just prison. In F. T. Cullen, C. L. Jonson & M. K. Stohr (Eds.), *The American prison: Imagining a different future* (pp. 111-126). Thousand Oaks, CA: Sage.

Holsinger, K. (2014). The feminist prison. In F. T. Cullen, C. L. Jonson & M. K. Stohr (Eds.), *The American prison: Imagining a different future* (pp. 87-110). Thousand Oaks, CA: Sage.

Imbens, G. W., & Kalyanaraman, K. (2009). *Optimal bandwidth choice for the regression discontinuity estimator*. NBER Working Paper 14726. National Bureau of Economic Research. Cambridge, MA.

Irwin, J. (2005). *The warehouse prison: Disposal of the new dangerous class*. Los Angeles: Roxbury

Jalbert, S. K., Rhodes, W., Flygare, C., & Kane, M. (2010). Testing Probation Outcomes in an Evidence-Based Practice Setting: Reduced Caseload Size and Intensive Supervision Effectiveness. *Journal of Offender Rehabilitation, 49*, 233-253.

King, G., & Roberts, M. E. (2015). How robust standard errors expose methodological problems they do not fix and what to do about it. *Political Analysis, 23*, 159-179.

Lee, D., & McCrary, J. (2009). *The deterrent effect of prison: Dynamic theory and evidence*. Unpublished manuscript, Department of Economics, Princeton University, Princeton, NJ.

Loeffler, C. E., & Grunwald, B. (2015). Processed as an adult: A regression discontinuity estimate of the crime effects of charging nontransfer juveniles as adults. *Journal of Research in Crime & Delinquency, 52*, 890-922.

Loughran, T. A., Mulvey, E. P., Schubert, C. A., Fagan, J., Piquero, A. R., & Losoya, S. H. (2009). Estimating a dose-response relationship between length of stay and future recidivism in serious juvenile offenders. *Criminology, 47*, 699-740.

Maddan, S., Miller, J. M., Walker, J. T., & Marshall, I. H. (2011). Utilizing Criminal History Information to Explore the Effect of Community Notification on Sex Offender Recidivism. *Justice Quarterly, 28*, 303-324.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics, 142*, 698-714.

Mears, D. P., Cochran, J. C., & Bales, W. D. (2012). Gender differences in the effects of prison on recidivism. *Journal of Criminal Justice, 40*, 370-378.

Mears, D. P., Cochran, J. C., & Cullen, F. T. (2014). Incarceration heterogeneity and its implications for assessing effectiveness of imprisonment on recidivism. *Criminal Justice Policy Review*.

Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social science*. New York: Oxford.

Nagin, D., Cullen, F. T., & Jonson, C. L. (2009). Imprisonment and Reoffending. In M. Tonry (Ed.), *Crime and Justice* (Vol. 38, pp. 115-200). Chicago: University of Chicago Press.

Nagin, D. S., & Snodgrass, M. G. (2013). The effect of incarceration on re-offending. *Journal of Quantitative Criminology, 29*, 601-642.

National Research Council. (2014). *The growth of incarceration in the United States: Exploring causes and consequences*. Washington, DC: The National Academies Press.

Paternoster, R. (1987). The deterrent effect of the perceived certainty and severity of punishment: A review of the evidence and issues. *Justice Quarterly, 4*, 173-217.

Paternoster, R., & Iovanni, L. (1989). The labeling perspective and delinquency: An elaboration of the theory and an assessment of the evidence. *Justice Quarterly, 6*, 359-394.

Rauma, D., & Berk, R. A. (1987). Remuneration and Recidivism: The Long-Term Impact of Unemployment Compensation on Ex-Offenders. *Journal of Quantitative Criminology, 3*, 3-27.

Savage, C. (2013, August 12, 2013). Dept. of Justice seeks to curtail stiff drug terms. *The New*

York Times, p. A1.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*: Cengage Learning.

Smith, P., Goggin, C., & Gendreau, P. (2002). *The Effects of Prison Sentences and Intermediate Sanctions on Recidivism: General Effects and Individual Differences*. Ottawa: Public Works and Government Services Canada.

Sykes, G. M. (1958). *The society of captives: A study of a maximum security prison*. Princeton, NJ: Princeton University Press.

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex-post facto experiment. *Journal of Educational Pscyhology, 51*, 309-317.

Tonry, M. (2004). *Thinking about crime: Sense and sensibility in American penal culture*. New York: Oxford University Press.

Villettaz, P., Killias, M., & Zoder, I. (2006). *The Effects of Custodial vs. Non-custodial Sentences on Re-offending; A Systematic Review of the State of the Knowledge* Oslo: Campbell Collaboration Crime and Justice Group.

Wacquant, L. (2001). Deadly Symbiosis: When Ghetto and Prison Meet and Mesh. *Punishment & Society, 3*, 95-133.

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

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Worrall, J. L., & Morris, R. G. (2011). Inmate Custody Levels and Prison Rule Violations. *Prison Journal, 91*, 131-157.

Table 1. Comparison between Offenders Below and Above Cut-Point

	Full Sample		Restricted Range (Within Optimal Bandwidth)	
	Below (<= 44) (n = 264,595)	Above (>44) (n = 66,376)	Just Below (n = 7,268)	Just Above (n = 2,392)
Variable	Mean (SD) or %	Mean (SD) or %	Mean (SD) or %	Mean (SD) or %
Age	31.83 (10.36)	32.36 (10.73)**	32.62 (9.97)	32.82 (10.11)
Race/Ethnicity		$\chi^2 = 2{,}000^{**}$		$\chi^2 = 2.78$
Black	41%	50%	51%	53%
Hispanic	9%	10%	8%	8%
Whites	50%	40%	41%	39%
Gender		$\chi^2 = 3700^{**}$		$\chi^2 = 0.41$
Male	78%	88%	87%	88%
Female	22%	12%	13%	12%
Offense Type		$\chi^2 = 3000^{**}$		$\chi^2 = 5.21$
Violent Off.	12%	37%	23%	23%
Property Off.	34%	27%	26%	25%
Drug Off.	39%	26%	36%	36%
Other Off.	15%	10%	15%	16%
Prior Conviction Points	.71 (1.82)	1.61 (3.37)**	1.55 (2.95)	1.50 (2.86)
Prior Prison Sentences	.26 (0.78)	.91 (1.50)**	.81 (1.35)	.93 (1.39)*1
Prior Supervision Sentences	1.56 (1.28)	1.95 (1.72)**	2.23 (1.72)	2.24 (1.77)
Prior Supervision Violations	.61 (1.07)	1.08 (1.43)**	1.28 (1.43)	1.40 (1.50)**1
Sentenced to Prison	4%	38%**	13%	39%**

$^{*}p < .05$; $^{**}p < .01$; [1] Difference is not statistically significant at .05 level after controlling for total sentence points.

Table 2: Effect of Scoring to Prison on Reconviction Probit Regression within Various Bandwidths: Coefficients (Robust SE)

Variable	All Cases	Blacks	Hispanics	Whites	Males	Females
50% Optimal Bandwidth						
Scored to Prison	.097 (.065)	.149 (.090)	-.112 (.231)	.096 (.105)	.151* (.070)	-.272 (.182)
Sentence Points	-.013 (.062)	-.055 (.087)	.178 (.223)	-.019 (.100)	-.044 (.067)	.191 (.182)
Constant	-.184** (.031)	-.023 (.043)	-.204 (.117)	-.400** (.050)	-.183** (.034)	-.209* (.088)
N	5,829	2,996	460	2,373	5,131	698
Estimated Difference in Probability of Reconviction[1]	.038 (.026)	.059 (.036)	-.043 (.088)	.036 (.040)	.060 (.028)	-.101 (.066)
100% Optimal Bandwidth						
Scored to Prison	.053 (.048)	.111# (.066)	-.247 (.175)	.058 (.077)	.098# (.051)	-.244# (.135)
Sentence Points	-.001 (.024)	-.036 (.033)	.163 (.088)	-.009 (.038)	-.023 (.026)	.131* (.067)
Constant	-.192** (.024)	-.025 (.033)	-.253** (.089)	-.412** (.039)	-.188** (.026)	-.235** (.069)
N	11,047	5,703	851	4,493	9,677	1,370
Estimated Difference in Probability of Reconviction	.020 (.019)	.044 (.026)	-.090 (.063)	.058 (.077)	.039 (.020)	-.090 (.049)
200% Optimal Bandwidth						
Scored to Prison	.021 (.034)	.054# (.048)	-.077 (.127)	.005 (.055)	.043 (.037)	-.131 (.099)
Sentence Points	.015 (.009)	.004 (.012)	.052 (.033)	.008 (.014)	.008 (.009)	.053* (.025)
Constant	-.188** (.018)	-.010 (.025)	-.292** (.066)	-.408** (.029)	-.171** (.019)	-.306** (.052)
N	22,094	11,294	1,639	9,161	19,221	2,873
Estimated Difference in Probability of Reconviction	.008 (.013)	.022 (.019)	-.029 (.047)	.002 (.020)	.017 (.014)	-.048 (.036)

Note: Sentence points has been centered at the cut score (i.e., sentence points – 44)

[1] The difference in predicted probability of reconviction for cases scored to prison minus the predicted probability of reconviction for cases not scored to prison.

#$p < .10$, *$p < .05$, **$p<.01$

Table 3. Effect of Imprisonment on Reconviction Instrumental Variable Analysis: Probit Coefficients (Robust SE)

Variable	All Cases	Blacks	Hispanics[1]	Whites	Males	Females
50% Optimal Bandwidth						
Scored to Prison	.433 (.286)	.584$^{\#}$ (.353)	-----	.433 (.463)	.658* (.302)	-1.341 (.897)
Sentence Points	-.020 (.066)	-.053 (.086)	-----	-.032 (.111)	-.057 (.071)	.174 (.173)
Constant	-.246** (.069)	-.101 (.087)	-----	-.465** (.113)	-.280** (.075)	-.069 (.173)
N	5,829	2,996	-----	2,373	5,131	698
Estimated Difference in Probability of Reconviction	.171 (.113)	.233 (.141)	-----	.162 (.173)	.260 (.120)	-.380 (.341)
100% Optimal Bandwidth						
Scored to Prison	.212 (.213)	.436$^{\#}$ (.263)	-----	.217 (.349)	.400$^{\#}$ (.222)	-1.333$^{\#}$ (.739)
Sentence Points	-.004 (.028)	-.041 (.036)	-----	-.010 (.046)	-.029 (.030)	.162* (.082)
Constant	-.221** (.054)	-.089 (.070)	-----	-.439** (.087)	-.246** (.057)	-.082 (.148)
N	11,047	5,703	-----	4,493	9,677	1,370
Estimated Difference in Probability of Reconviction	.083 (.084)	.174 (.105)	-----	.080 (.129)	.158 (.088)	-.378 (.277)
200% Optimal Bandwidth						
Scored to Prison	.071 (.144)	.199 (.178)	-.573 (1.129)	-.035 (.246)	.164 (.151)	-.575 (.455)
Sentence Points	.014 (.011)	.003 (.014)	.060 (.053)	.011 (.019)	.006 (.011)	.063* (.033)
Constant	-.196** (.038)	-.039 (.050)	-.235 (.184)	-.397** (.064)	-.194** (.040)	-.237 (.104)
N	22,094	11,294	1,639	9,161	19,221	2,873
Estimated Difference in Probability of Reconviction	.028 (.056)	.079 (.071)	-.215 (.424)	-.013 (.090)	.065 (.060)	-.213 (.169)

Note: Instrumental variable analyses use sentence assigned by the CPC (i.e., scored to prison or not) as an instrumental variable to identify exogenous variation in actual imprisonment decisions. Sentence points has been centered at the cut score.

[1] In the Hispanic specific models using 50% and 100% optimal bandwidth, the instrumental variable did not predict the endogenous variable; therefore, the results of these models did not meet the requirements of instrumental variable analysis.

$^{\#}p < .10$, $^{*}p < .05$, $^{**}p<.01$

Figure 1. Probability of Prison Sentence by Sentencing Guideline Points

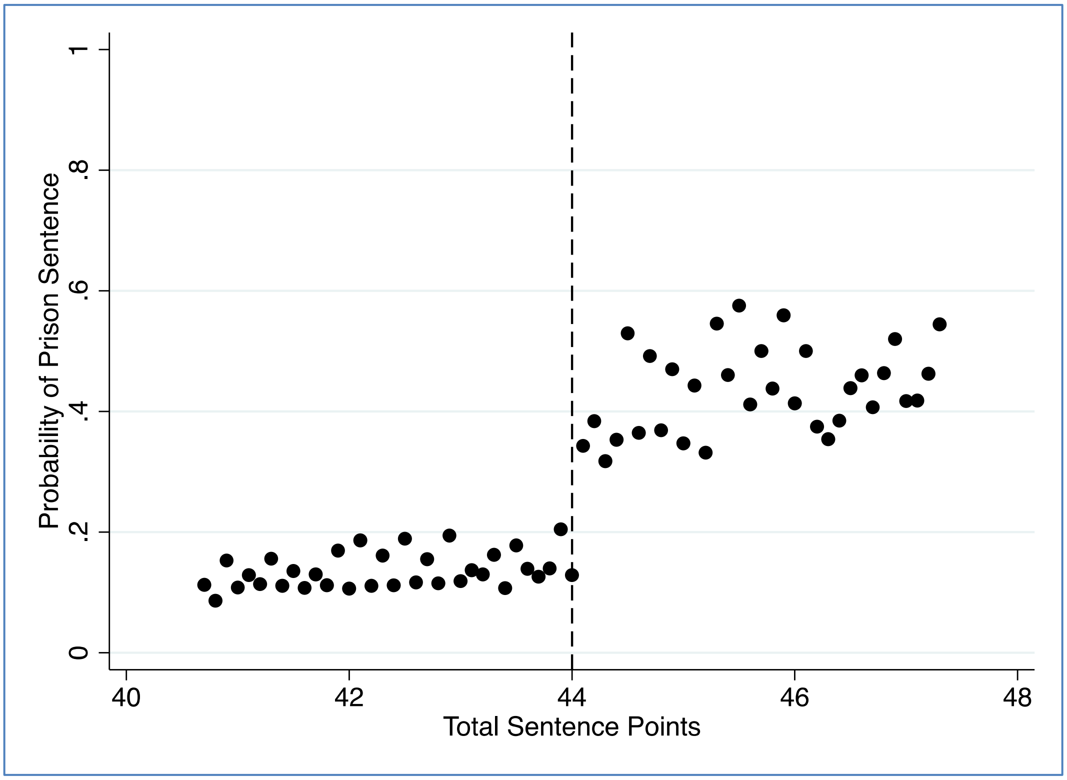

Figure 2. Kernel Density of Sentencing Guideline Points

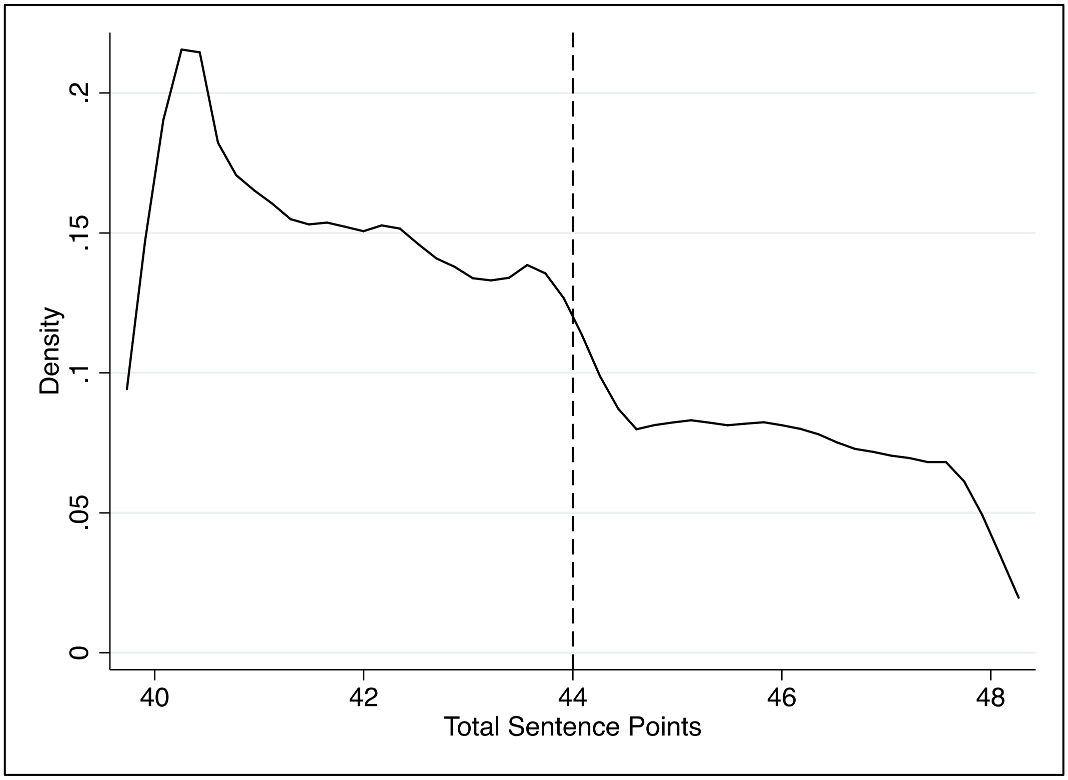

Figure 3. Probability of Reconviction by Sentence Points: Restricted Range