A Revolving Door?

A Meta-Analysis of the Impact of Custodial Sanctions on Reoffending

Damon M. Petrich
University of Cincinnati

Travis. C. Pratt
University of Cincinnati Corrections Institute
Harris County Community Supervision and Corrections Department

Cheryl Lero Jonson
Xavier University

Francis T. Cullen
University of Cincinnati

Contact:

Damon M. Petrich, School of Criminal Justice, University of Cincinnati, Cincinnati, OH 45226-0389. Email: petricdm@ucmail.uc.edu

Recommended Citation:

Petrich, D. M., Pratt, T. C., Jonson, C. L., & Cullen, F. T. (2020). A Revolving Door? A Meta-Analysis of the Impact of Custodial Sanctions on Reoffending. Working paper, University of Cincinnati.

A Revolving Door? A Meta-Analysis of the Impact of Custodial Sanctions on Reoffending

Abstract

Research Summary: Despite a downsizing of the prison population over the last decade, the United States remains the world's largest incarcerator. While prior meta-analytic reviews have indicated prisons have a null or slight criminogenic effect, there has been a substantial increase in primary studies assessing the impact of custodial sanctions on reoffending since 2010. The current study conducts the most comprehensive and rigorous meta-analytic review of research to date comparing the effects of custodial and non-custodial sanctions on reoffending. Using a sample of 959 effect sizes calculated from 116 studies, this review builds on prior research by (a) conducting the first inclusive review of the research in over a decade; (b) drawing on advancements in modeling techniques that could account for the precision and statistical dependence of effect sizes; and (c) examining whether a wide variety of methodological characteristics moderate mean effect sizes. The results of multilevel analyses revealed that custodial sanctions have a weak criminogenic effect on reoffending and that this effect is relatively robust across a wide variety of methodological moderators.

Policy Implications: America's reliance on incarceration is unlikely to wane in the near future. However, the results of the current study illustrate that incarceration is not achieving one of its oft-desired goals of deterrence. While there is certainly a subset of high-risk offenders for whom incapacitation is warranted, in general, placing individuals in custodial sanctions appears to contribute to, rather than reduce, reoffending. As a result, sentencing low-level offenders and marginal felons to non-custodial sanctions may lead to considerable savings for correctional departments and allow offenders to be connected to and contribute to their communities while reducing reoffending. For the sake of public safety and those for whom incarceration is appropriate, however, more research is needed on the in-prison social and psychological dynamics increase post-release criminal behavior.

After a nearly 600% increase in the rate of imprisonment in the United States between 1972 and 2007 (Nellis & King, 2009; Pager, 2007), this trend is slowly beginning to reverse, with rates declining by an average of 1.2% per year between 2007 and 2016 (Bureau of Justice Statistics, 2018). Still, the latest nationwide data released by the Bureau of Justice Statistics (2018) show that, as of December 2016, 2,162,400 individuals were incarcerated in state/federal prisons or local jails. Even with declines over the past decade, the United States maintains its dubious title as the world leader in imprisonment with approximately 655 prisoners per 100,000 citizens, followed by El Salvador (604) and Turkmenistan (526) (World Prison Brief, 2018). The imprisonment rates of other culturally comparable countries, such as Canada (114), the United Kingdom (140), and Australia (172), are but a fraction of America's (World Prison Brief, 2018). Taken together, government agencies in the United States spend approximately \$80 billion per year on corrections (National Academy of Sciences, 2014), with average incarceration costs of around \$30,000 per inmate, per year (PEW Center on the States, 2009; Vera Institute, 2017).

Prior commentaries (e.g., Clear & Frost, 2014; Garland, 2001; Pfaff, 2017) illustrate that the meteoric growth in the use of custodial sanctions has been justified on the grounds that locking offenders up provides a uniquely effective means of deterring future criminal behavior. Indeed, if humans are viewed as rational actors who aim to seek pleasure and avoid pain, and, accordingly, weigh the costs and benefits of their actions (Beccaria, 1764 [1983]; Becker, 1968), "common sense" suggests that the loss of freedoms imparted by a stint of incarceration would be a more effective deterrent than other, non-custodial sanctions such as probation or mandated participation in non-residential treatment (Pratt, Cullen, Blevins, Daigle, & Madensen, 2006). An alternative perspective, however, is that incarceration has iatrogenic rather than deterrent effects on individual offenders (Blevins, Listwan, Cullen, & Jonson, 2010; Cid, 2009; Nagin, Cullen, &

Jonson, 2009). Close-quarters cohabitation with and learning from other inmates; the strains of increased exposure to violence and the loss of one's freedoms; the fraying of ties to one's support networks; and collateral consequences such as barriers to employment and housing have all been suggested as factors that may increase reoffending among those sentenced to serve time in custody (Bayer, Hjalmarsson, & Pozen, 2009; Listwan, Sullivan, Agnew, Cullen, & Colvin, 2013; Sampson & Laub, 1993).

In adjudicating between these two perspectives, results from individual studies are mixed, with some indicating a deterrent effect of incarceration (e.g., Hjalmarsson, 2009; Jones & Ross, 1997), others finding null effects (e.g., Aarten & Denkers, 2015; Gaes, Bales, & Scaggs, 2016), and still others showing criminogenic effects (e.g., Mears & Cochran, 2018; Nieuwbeerta, Nagin, & Blokland, 2009). The reasons for this heterogeneity are myriad; individual studies vary widely in the characteristics of their samples, the analytical methods employed, the confounding influences accounted for, and the type of reoffending outcomes examined. However, by aggregating the findings of a body of research and explicitly accounting for the heterogeneity of effects, the technique of meta-analysis allows for the estimation of a mean effect for a given treatment on an outcome of interest (Borenstein, Hedges, Higgins, & Rothstein, 2009; Wolf, 1986). To date, three such meta-analyses have been conducted to determine the impact of custodial sanctions on reoffending when compared to noncustodial sanctions, each finding either a null or criminogenic effect (Jonson, 2010; Smith, Goggin, & Gendreau, 2002; Villettaz, Gillieron, & Killias, 2015).

With the recent downsizing of the prison population in the United States, we present the results of an updated meta-analysis of the literature comparing the reoffending outcomes of offenders given custodial and noncustodial sanctions (Ghandnoosh, 2019). Our analysis builds

upon previous meta-analytic work in three important ways. First, through the inclusion of a broader range of research produced prior to Villettaz et al. (2015) and a large number of studies since, the present study represents the first comprehensive meta-analytic review of the literature comparing custodial to non-custodial sanctions in over a decade. Second, given the wide variability in the samples and methodologies employed in this body of work, we also assessed moderators of effect size estimates. Prior meta-analytic work has either not included such moderator analyses at all (Villettaz et al., 2015) or examined only a handful of potential moderators (Smith et al., 2002). In the current study, we look at a broad range of characteristics (e.g., age and gender composition, type of sanctions served, whether theoretically relevant constructs are statistically controlled), allowing us to determine how robust effect sizes are to differences in samples and methodological choices made, and to provide guidance for future research on the topic. Third, our analyses use a multilevel modeling (MLM) approach (Hox, 2017), allowing (a) individual effect sizes to be weighted by their precision rather than sample size and (b) the inclusion of multiple effect sizes from within individual studies while accounting for their statistical dependence; as discussed further below, both of these issues have been unsatisfactorily addressed in prior meta-analyses of the sanctions literature. However, to set the context of our inquiry, we first discuss the state of the evidence for the effects of incarceration on reoffending.

Estimating the Impact of Custodial Sanctions on Reoffending

Sources of Heterogeneity

As noted at the outset of the present article, there are two conflicting perspectives on the general impacts of incarceration on reoffending¹. The deterrence perspective suggests that because the incarceration experience is aversive and more severe than non-custodial sanctions, it should deter the future criminal behavior of individual offenders. The criminogenic experience perspective, on the other hand, suggests that strenuous experiences, social learning, and official labeling make reoffending more likely for those incarcerated than if they had been sentenced to alternative, non-custodial sanctions. Findings from prior research that have attempted to adjudicate between these two hypotheses have been mixed. There are myriad reasons for the observed heterogeneity in this body of research; we discuss three here that are of most relevance.

First, there are wide variations in the demographic and criminal history-related characteristics of the samples examined. Some studies focus explicitly on a single type of offender, looking, for example, solely at the effects of custodial versus non-custodial sanctions on offenders convicted of drug (Gibbs, Lytle, & Wakefield, 2019), domestic violence (Trevena & Poynton, 2016), or sex offenses (Nunes, Firestone, Wexler, Jensen, & Bradford, 2007). Other studies examine sanction effects exclusively amongst males (Cochran, Mears, & Bales, 2014), females (Hedderman & Jolliffe, 2015), or stratify analyses by gender (Mitchell, Cochran, Mears, & Bales, 2017). And still other studies analyze only or stratify analyses according to whether samples comprise first-time or repeat offenders (Kraus, 1974; Mears & Cochran, 2018; Nieuwbeerta et al., 2009). Most studies, however, include offenders with a range of such characteristics, and between-sample variations in the distributions of gender, age, criminal history, and the like can be quite large.

¹ Given that these two perspectives are well-known and have been reviewed extensively elsewhere, interested readers are referred to the works of Chalfin and McCrary (2017), Cullen, Jonson, and Nagin (2010), Jonson (2010), and Nagin et al. (2009) for detailed discussions.

Second, as observed by both Villettaz et al. (2015) and Nagin et al. (2009), there is also considerable heterogeneity in the types and lengths of sanctions—both custodial and non-custodial—that are compared in the extant literature. In our meta-analytic review, types of custodial sanctions include prison, jail, juvenile detention, residential treatment, boot camp, and shock incarceration, among others. On the non-custodial side, these include regular probation, intensive probation, electronic monitoring, community service/fines, fully suspended sentences, dismissals, and non-residential treatment. Beyond the types of sanctions examined, there is also variability in combinations that are compared, the duration of sentences served, and the types of individuals serving them. For example, both Gillman, Hill, and Hawkins (2015) and Wermink, Blokland, Nieuwbeerta, Nagin, and Tollenaar (2010) examine the impact of short-term incarceration (i.e., less than 60 days) on reoffending. However, Gillman et al. do so with a sample of juveniles and compare against those arrested but not incarcerated, while Wermink et al. compare short-term custody with community service within a sample of adults. Some analysts working with large datasets elect to collapse multiple types of sanctions into singular “custodial” and “non-custodial” categories. In their analyses of 330,971 offenders in Florida, for instance, Mitchell et al. (2017) combined those who had served time in prison with jail inmates, and those who had served regular probation with those on intensive probation. Working with these same data, Mears and Cochran (2018) later chose to disaggregate these groupings for increased specificity in effect estimates.

A final factor that likely accounts for much of the variability in findings from prior research is large discrepancies in the methodologies employed. Although the randomized controlled trial (RCT) is certainly the gold standard for reaching estimates of the causal effect of interventions (Shadish, Cook, & Campbell, 2002), the practical and ethical realities of randomly

assigning offenders to prison versus probation make RCT-based evaluations exceptionally rare. Only five such studies have been conducted to date (Barton & Butts, 1990; Bergman, 1976; Killias, Aebi, & Ribeaud, 2000; Killias, Gilliéron, Villard, & Poggia, 2010; Schneider, 1986). Among them, the Killias et al. (2000, 2010) studies rely on the same data and, together with the Schneider (1986) study, examine the impact of two weeks or less of incarceration. The Barton and Butts (1990) RCT examines a longer period of incarceration (avg = 12.8 months), while the amount of time is unspecified in Bergman (2000); both use data that are over 30 years old.

In light of the difficulties with conducting RCTs, the vast majority of prior research has used observational data coupled with widely varying levels of analytical rigor to determine the effects of custodial sanctions. At one end of the spectrum are those analyses which employ neither matching procedures nor statistical controls to account for non-random treatment assignment. This is often the case within government reports for state and federal departments of correction; for example, a recent report by McAlister, Officer, and Sanchagrin (2019) provided the three-year recidivism rates for bi-annual cohorts of offenders released from prison or placed on probation in Oregon between 1998 and 2016. Not surprisingly, the majority of effect size estimates from this and similar reports point to a large criminogenic effect of prison, likely due to the fact that, on average, individuals sentenced to prison differ substantially from those sentenced to probation (Aizer & Doyle, 2015; Sweeten & Apel, 2007; Wermink et al., 2010). Other studies improve estimates by incorporating regression-based techniques for covariate adjustment or quasi-experimental methods such as exact matching, propensity score matching (PSM), or instrumental variable analysis. Nagin et al. (2009; see also Bales & Piquero, 2012; Gaes et al., 2016) have previously reviewed the problems associated with regression-based analyses of prison effects (e.g., imprecision of specifying age effects, linearity assumption, over-

parameterization), urging researchers to instead opt for quasi-experimental designs. Even among these methodologies, however, there is much variability in model specification that can be consequential for effect size estimates. Bales and Piquero (2012) show, unsurprisingly, that the sizes of estimates from both exact matching and PSM are sensitive to which theoretical constructs matching procedures account for. Similarly, Gaes et al. (2016) show that, despite matching on the same set of covariates, radius, coarsened exact, and exact matching techniques yield substantially different estimates in terms of size, directionality, and statistical significance.

Prior Meta-Analytic Reviews of the Literature

By aggregating the results of an entire body of work, the technique of meta-analysis is advantageous in that it allows for the estimation of an overall mean effect size. Further, moderation analyses allow for the examination of how well between- and within-study differences in the characteristics reviewed above (i.e., samples, the types and lengths of sanctions served, and methodological considerations) explain the heterogeneity of effect size estimates (Borenstein et al., 2009; Pratt & Cullen, 2000; Wolf, 1986). To date, three such analyses have probed whether custodial sanctions have a deterrent or criminogenic effect when compared to non-custodial sanctions. All three meta-analytic reviews found that custodial sanctions either have no effect or lead to slight increases in reoffending.

First, in a report to the Solicitor General of Canada, Smith et al. (2002) investigated this issue by drawing on 104 effect sizes garnered from 31 studies. The mean phi coefficient in their analyses was between .00 and .03 (unweighted and weighted by sample size, respectively), indicating a null or small criminogenic effect of custodial sanctions on reoffending. Smith et al. examined only a small number of potential moderators; however, such analyses indicated effect sizes were larger in studies of juveniles than adults ($\phi = .08$ vs. $.03$) and in studies of higher

compared to lower methodological quality ($\phi = .10$ vs. $-.01$). The modeling strategy used by Smith et al. (2002) did not explicitly account for the statistical dependence of multiple effect sizes drawn from some individual studies. However, sensitivity analyses showed that the mean effect size estimate remained the same when these multiple, overlapping effect sizes were excluded.

Second, Jonson (2010) meta-analyzed 177 effect size estimates from 57 studies, finding a small criminogenic effect of custodial sanctions (mean $r = .144$). In contrast to Smith et al. (2002), Jonson's (2010) moderation analyses showed that effect sizes were larger in samples comprised exclusively of adults and that methodological quality was not a significant moderator. Further, effect size estimates varied by the types of both custodial and non-custodial sanctions served and the gender composition of the samples examined. Although the larger number of effect sizes and methodological moderators examined provided an improvement over Smith et al.'s (2002) work, Jonson's (2010) modeling technique did not take into account the statistical dependence of effect sizes, nor were sensitivity analyses conducted to examine the influence of such dependence on mean effect size estimates.

Third, in the most recent meta-analytic review, Villettaz et al. (2015) provided an update to an earlier Campbell Review (Villettaz, Killias, & Zoder, 2006) by examining sanction effects among 14 studies using random effects regression techniques. Looking solely at the results from RCTs and natural experiments ($k = 5$), custodial sanctions had a null effect on reconviction relative to non-custodial sanctions (mean odds-ratio [OR] = .946, $p > .05$). Among quasi-experimental studies ($k = 9$), a small but criminogenic effect of custodial sanctions was observed (mean OR = .684, $p < .001$). A limitation of this work is that Villettaz et al. (2015, p. 42) made the choice to include only a single effect size from each study, noting that "given the fact that the

overall results favoured the null hypothesis, the strongest effect sizes have been used as a conservative way to minimize the chance of obtaining a non-significant outcome.” Given the restrictive inclusion criteria adopted by Villettaz et al. (2015), the selection of one effect size per study, and the resultant small sample of studies from which to draw information, no moderator analyses were conducted.

The Current Study

Against the theoretical, empirical, and methodological backdrop laid out thus far, the aim of the current study was to conduct an updated meta-analytic review of research comparing the effects of custodial and non-custodial sanctions on reoffending. As discussed previously, these analyses expand upon prior work in three important ways. First, this study marks the first *comprehensive* meta-analytic review of the literature in over a decade (i.e., since Jonson, 2010). Although the review by Villettaz et al. (2015) was published five years ago, their analyses were limited by a restrictive set of inclusion criteria and the selection of only one effect size from each of the 14 studies included (e.g., one from among 36 effect sizes reported in Bales & Piquero, 2012). In addition to including 22 quasi-experimental studies published since Villettaz et al. (2015), the current study we relaxed the inclusion criteria to allow regression-based and unadjusted comparisons of individuals sentenced to custodial and non-custodial sanctions. Doing so allowed us to take stock of the entire body of research, comprising 959 effect size estimates drawn from 116 studies. Of these, approximately two-thirds of our effect sizes (636) were drawn from 54 studies released since 2010.

A second and related contribution of the current study is that the large sample of effect size estimates and detailed coding of studies enabled the examination of putative moderators of effects. To be sure, including studies with wide variability in methodological rigor increases the

heterogeneity of effect sizes. Some have suggested that analyzing studies with such differences is akin to comparing apples and oranges (e.g., Eysenck, 1984; Sharpe, 1997), and that their inclusion in a meta-analysis may bias mean effect size estimates. However, we follow others (e.g., Borenstein et al., 2009; Smith, Glass, & Miller, 1980; Turanovic & Pratt, 2020) in arguing that information about both apples and oranges provides knowledge about fruit more broadly; similarly, information from both high- and low-quality studies provides knowledge about sanction effects and the impact of methodological rigor on estimates of effect size. To that end, we explicitly coded for a large number of characteristics such as within- and between-study differences in sample demographics (e.g., gender and age composition), the types of sanctions served (e.g., jail, prison, probation, intensive probation), variables controlled for/matched on (e.g., age, sentence length, prior record), and type of reoffending measured (e.g., rearrest, reincarceration). Doing so allowed for the explicit identification of factors which moderate effect sizes and should be considered in the design of future analyses of sanction effects.

Third, the current study capitalizes on improvements made in meta-analytic modeling techniques since the studies by Smith et al. (2002) and Jonson (2010), namely, multilevel modeling (MLM; see Borenstein et al., 2009; Hox, 2017). One advantage of the MLM approach is that greater weight is assigned to effect size estimates that are more reliable by explicitly accounting for their standard errors (discussed further in the Method section). This results in greater precision when estimating mean effect sizes than earlier techniques, such as weighting effect sizes solely by sample size (i.e., the practice used in both Smith et al., 2002 and Jonson, 2010; see Pratt et al., 2014 for further elaboration). The MLM approach also allows for the inclusion of multiple effect sizes from each study by accounting for the dependence of observations. Prior meta-analytic reviews conducted without MLM have not accounted for such

dependence (Jonson, 2010; Smith et al., 2002). Further, although Villettaz et al. (2015) did use MLM for their analyses, as noted above, they chose a single effect size from each study to include in their models. Both the failure to account for the dependence of observations within studies and the researchers' decisions to include specific effect sizes can bias meta-analytic results (Becker, 2000; Pratt & Cullen, 2000); the approach used in the present study alleviates both concerns, providing the most comprehensive estimate of the effect of custodial sanctions on recidivism to date.

Method

Sample

Studies published through May of 2019 were systematically gathered in four ways. First, all prior issues of several criminology/criminal justice journals were reviewed for studies which had compared the recidivism outcomes of groups sentenced to custodial versus non-custodial sanctions². Second, extensive searches of electronic databases (i.e., ProQuest Criminal Justice; EBSCO Criminal Justice Abstracts with Full Text; PsycINFO; Google Scholar; JSTOR Economics) were conducted with combinations of search terms that captured both the sanction type (i.e., “prison,” “imprison*,” “incarcerat*,” “boot camp,” and “custodial”) and outcomes of interest (i.e., “recidivism,” “reoffend*,” “rearrest,” “reincarcerate*,” and “reconvict*”). Third, the reference lists of studies from the first two steps and from prior meta-analytic reviews (i.e.,

² These journals included: *The Australian and New Zealand Journal of Criminology*, *British Journal of Criminology*, *Crime & Delinquency*, *Criminal Justice and Behavior*, *Criminology*, *Criminology & Criminal Justice*, *Criminology & Public Policy*, *European Journal of Criminology*, *International Journal of Offender Therapy and Comparative Criminology*, *Journal of Contemporary Criminal Justice*, *Journal of Criminal Justice*, *Journal of Developmental and Life-Course Criminology*, *Journal of Experimental Criminology*, *Journal of Offender Rehabilitation*, *Journal of Quantitative Criminology*, *Journal of Research in Crime and Delinquency*, *Justice Quarterly*, *Nordic Journal of Criminology*, *Punishment and Society*, *The Prison Journal*, *Youth Violence and Juvenile Justice*.

Jonson, 2010; Smith et al., 2002; Villettaz et al., 2015) were examined for studies not already captured. Fourth, state and federal correctional agencies' websites were searched for unpublished comparisons of post-sanction outcomes for custodial and non-custodial groups.

In reviewing the results of these searches of the extant research, the satisfaction of three criteria was required for each study's inclusion in the meta-analysis. First, the study must have included a group of offenders sentenced to spend time in a custodial setting and a comparison group given an alternative, non-custodial sanction. Second, the study must have included some measure of post-sanction criminal behavior (e.g., rearrest, technical violation). Third, the study must also have included sufficient information to calculate the common effect size used in the current analyses. These inclusion criteria are quite broad, although intentionally so; as described in further detail below, this allowed us to code for between- and within-study variation in a variety of sample and methodological characteristics (e.g., group equivalence, demographics, follow-up length) that potentially moderate effect sizes.

Through the literature search and evaluation of whether gathered studies met the inclusion criteria, the analytic sample comprises 959 effect sizes calculated from 116 studies that represent approximately 4.5 million individual offenders in 16 different countries³. The number of effect sizes exceeds the number of studies because the majority of studies included more than one statistical model from which an effect size could be calculated⁴. Including multiple effect

³ This figure is approximate given that some studies report the outcomes of multiple release cohorts that may comprise some of the same individuals.

⁴ As an example, Bales and Piquero's (2012) comparison of groups sentenced to prison and intensive probation (a) measured new felony convictions at one, two, and three years; (b) included models that used six increasingly stringent groups of statistical controls; and (c) modeled the effects of sanctions on reconviction, net of controls, using logistic regression, precision matching, and propensity score matching. Taken together, the Bales and Piquero study alone thus yielded 36 effect size estimates.

sizes from individual studies was done for two reasons. First, as pointed out elsewhere (e.g., Becker, 2000; Pratt & Cullen, 2000; Weisz et al., 2017), choosing a single effect size or averaging across effect sizes within a study can introduce researcher (e.g., picking the most showing the greatest criminogenic effect of custodial sanctions) and/or statistical bias (e.g., artificially reducing between-effect variance). Second, doing so also results in the loss of valuable within- and between-study information that can be used to explore the methodological and sample characteristics that may account for variance in effect sizes. Although including multiple effect sizes per study raises concerns over violating the assumption of statistical independence, as described in detail below, the multilevel modeling approach used in the current study adjusts for such dependence.

Effect Size Estimate

The effect size estimate in the current study represents the magnitude of the association between receiving a custodial (as compared to a non-custodial) sanction and subsequent recidivism. Effect sizes were calculated using the standardized correlation coefficient (i.e., r). The r coefficient was chosen because (a) its interpretation is generally more intuitive than other test statistics typically employed as meta-analytic effect size estimates (e.g., Cohen's d) and (b) other test statistics are easily converted into r -values⁵. Given the coding of sanction group-membership (i.e., 0 = non-custodial; 1 = custodial), positive r coefficients signify that being sentenced to a custodial sanction is associated with an increased likelihood of recidivism relative to those given a non-custodial sanction. For analytical purposes effect size estimates were

⁵ Specifically, t -ratios from linear models (e.g., ordinary least squares) can be converted using $r = t / \sqrt{t^2 + n - 2}$, while z -ratios from non-linear models (e.g., logistic and Cox regression) are converted using $r = z / \sqrt{z^2 + n}$ (see Rosenthal, 1994; Vartanian, Schwartz, & Bronwell, 2007). Odds-ratios are converted by taking their natural logarithm, converting to Cohen's d , and then converting to r (see Borenstein et al., 2009).

converted to $z(r)$ scores using Fisher's r to z transformation (i.e., $z = 0.5 * \ln(1 + r/1 - r)$; see Borenstein et al., 2009). As Pratt et al. (2014; see also Blalock, 1972) note, the sampling distribution of the $z(r)$ score is assumed to approach normality, while the distribution of r is not. For the multilevel linear modeling approach that we use, a normal distribution of effect size estimates is required for unbiased tests of statistical significance, accurate mean effect size estimates, and tests of their moderation.

We recognize concerns over the potential drawbacks of effect sizes drawn from both bivariate and multivariate models. The main concern with bivariate models stems from failing to account for confounds, while in multivariate models, study-to-study differences in the confounds that *are* accounted for may be large (see discussion in Pratt et al., 2014). These issues are addressed in two ways. First, we code for whether effect sizes were drawn from bivariate or multivariate equations and separate out calculations of mean effect sizes accordingly. Second, we also code for variations in the characteristics of the sample and methodologies employed by researchers in the original study. These potential moderators are discussed next.

Moderators of Effect Size

Study characteristics. All moderators in this category were coded as categorical variables. *Author affiliation* captures whether listed authors of each study were affiliated with a university (48.91%), state/provincial agencies (37.85%), federal agencies (9.49%), a mix of affiliations (3.23%), or had some other organizational affiliation (0.52%). *Funding source* indicates whether authors reported receiving financial support from state/provincial agencies (37.64%), federal agencies (19.77%), private organizations (e.g., non-profits; 1.77%), some other funding source (3.96%), no funding (3.55%), or if the receipt of funding was unreported (34.31%). *Publication decade* taps the decade during which the study was published (1960s =

1.98%; 1970s = 0.52%; 1980s = 3.75%; 1990s = 14.18%; 2000s = 13.24%; 2010s = 66.32%).

Study location refers to the country from which the sample was drawn, including the United States (78.31%), Canada (1.25%), the United Kingdom (5.32%), Australia (3.44%), Nordic countries (i.e., Denmark, Finland, and Norway; 4.59%), the Netherlands (3.75%), or some other country (3.34%). Finally, we also coded for *publication type*, including peer-reviewed journal articles (49.95%), state/local reports (34.83%), federal reports (5.94%), theses/dissertations (1.67%), book chapters (1.04%), or any other type of document (e.g., unpublished working papers; 6.57%).

Sample characteristics. Characteristics of the samples compared in calculations of each effect size were also coded, all of which as categorical variables. For *age composition*, most samples were exclusively comprised of adults (66.25%), followed by samples which were exclusively juvenile (16.19%), more than 80% adult (2.73%), or more than 80% juvenile (0.53%). A further 14.30% of effect sizes came from studies which did not report the age composition of their samples. For *gender composition*, the majority of samples were either exclusively male (15.85%) or comprised of more than 80% males (33.47%), followed by those which were a “mixed group” (i.e., less than 80% male; 13.35%), or exclusively female (3.13%). A further 34.20% of effect sizes were derived from studies in which gender composition was unreported.

The types of non-custodial and custodial sentences served by samples were coded as well. In terms of *non-custodial sanction type*, the majority of effect sizes were based on samples sentenced to probation (54.01%), followed by intensive probation (15.22%), community service/fines (7.51%), suspended sentences or dismissals (5.94%), electronic monitoring/house arrest (3.13%), or an explicitly treatment-focused non-custodial sanction (e.g., community-based

drug treatment; 1.67%). Approximately 13% of effect sizes were calculated from studies in which the type of non-custodial sanction was categorized as “Other,” capturing both unspecified non-custodial sanctions and low-frequency sanctions (e.g., restorative justice). For the custodial sanction type variable, the majority of effect sizes came from studies in which the custodial group was sentenced to prison (63.71%), followed by juvenile detention (10.84%), jail (8.65%), boot camp/shock incarceration (3.75%), or a secure residential facility (0.94%). A further 12.10% were coded as “Other,” which indicates unspecified or mixed sanction types (e.g., where the custodial sample was sentenced to either prison or jail). Although the majority of effect sizes were drawn from studies which did not report the mean length of time spent in custody (74.55%), for those which did, the *time in custody* moderator variable reflects that mean length of time. Most of the incarcerated groups served between one and less than six months in custody (53.28%), followed by those sentenced to between 6 months and less than one year (22.13%), two years or more (10.66%), less than one month (10.25%), or between one and less than two years (3.69%).

Research design. Several aspects of research design and model specification were also coded for each effect size included in the current meta-analysis. *Recidivism measure* captures the type of recidivism that custodial and non-custodial groups were compared on. These include new convictions (42.65%), arrests/charges (30.14%), reincarceration (22.63%), technical violations (1.56%), mixed measures (0.31%), and other types of recidivism (e.g., self-reported offending; 2.71%). *Length of follow-up* reflects the period of time samples were tracked for subsequent recidivism incidents: one year or less (25.29%), more than one to two years (21.51%), more than two to three years (35.78%), more than three to four years (4.30%), or greater than four years (13.12%).

Finally, for each effect size, dichotomous indicators for a wide range of possible confounds were coded, with scores of 1 indicating the focal characteristic was either matched on prior to analysis or statistically controlled for in the analysis, and scores of 0 indicating the focal characteristics was neither matched on nor controlled for. The characteristics coded for included age, gender, marital status, employment status, education level, socioeconomic status, race/ethnicity, type of conviction offense, age at first offense, prior record, substance abuse, mental health problems, risk level, and length of sentence.

Analytic Plan

We use the MLM procedures described by Hox (2017) and employed in other recent meta-analyses (e.g., Myers, Turanovic, Lloyd, & Pratt, 2020; Pratt et al., 2014; Pyrooz, Turanovic, Decker, & Wu, 2016) to estimate the mean effects of custodial versus non-custodial sanctions on reoffending. A two-level MLM framework is appropriate in the present study given that effect size estimates are nested; that is, Level 1 of the data corresponds to individual statistical models producing each effect size ($N = 959$), while Level 2 corresponds to the studies from which (often multiple) effect sizes were drawn ($k = 116$). At Level 1, then, effect sizes within studies often share the same samples and methods of data collection, while assessing different outcomes and with different model specifications. Without accounting for this hierarchical nature of the data, the assumptions of the statistical independence of observations and uncorrelated error are violated (Snijders & Bosker, 2012), thus increasing the likelihood of biased tests of statistical significance because of artificially deflated standard errors and narrow confidence intervals (Kreft & de Leeuw, 1998). However, the MLM framework used here resolves these issues and accounts for both within- and between-study sources of dependence through the inclusion of a unique random effect for each organizational unit (see Pratt et al.,

2014). Calculation of the intraclass correlation coefficient for the full analytic sample in our data indeed provide further evidence of the necessity to account for both sources of variation; 49.22% of the variance in effect size estimates was within studies ($\sigma^2 = .0062$, $p < .001$), while 50.78% was between studies ($\sigma^2 = .0065$, $p < .001$).

Another issue with hierarchical meta-analytic data is that a portion of the variance of Level 1 effect size estimates is assumed to be known (Hox & de Leeuw, 2003). Explicitly accounting for this variance is crucial given that effect size estimates are drawn from other studies that vary in precision (e.g., because of differences in sample size, model specification). To do so, we calculated standard errors for each effect size using $\sigma = \sqrt{1/(n - 3)}$ for bivariate effect size estimates and $\sigma = \text{Fisher's } z(r)/(b/SE)$ for multivariate effect size estimates (Lipsey & Wilson, 2001). These effect size standard errors were incorporated into the random part of the Level 1 equation, while constraining the variance to one (Hox, 2017; see also Myers et al., 2020; Pratt et al., 2014; Pyrooz et al., 2016), thus enabling models to account for both the known and within-study variance of effect size estimates.

With these considerations in mind, the analyses presented below proceeded in three stages. First, we estimated the overall mean effect of custodial sanctions (as compared to non-custodial sanctions) on recidivism. We examined these mean effect sizes across the entire sample of 959 effect size estimates, as well as when looking only at estimates derived from bivariate or multivariate models, and those derived from odds-ratios⁶. In the second phase, we conducted a series of bivariate moderator analyses to determine whether the overall mean effect of custodial sanctions on recidivism was robust across variations in characteristics of the samples studied and

⁶ We specifically examined effect sizes derived from odds-ratios, which can be notoriously unstable (and typically overestimated) when the split on the distribution of a dichotomous outcome departs from 50/50% (Persoskie & Ferrer, 2017).

the methodologies used to do so. All bivariate moderator analyses were conducted by entering only the focal moderator variable into the regression equation. Third, we examined the influence of significant moderators from the previous step when considered together in a multivariate meta-regression model, both with the full sample and restricted to effect sizes from multivariate models. All analyses were conducted in Stata 15 using the *-meglm-* command with maximum likelihood estimation.

Results

[insert Table 1 about here]

Strength of Effects

The overall mean effect size estimates (Mz) for the association between custodial sanctions and recidivism are reported in Table 1. These estimates were drawn from variance-known models with the number of effects per study and sample size entered as covariates⁷. These two covariates were grand mean centered; thus, the reported Mz in each model should be interpreted as the expected value of Mz when the number of effect sizes per and sample size are held at their respective means. As Table 1 shows, the overall Mz when looking at the entire sample of 959 effect size estimates was .080 ($p < .001$). Substantively, this effect size means that, on average, custodial sanctions have a positive but statistically significant effect on reoffending when compared to non-custodial sanctions, although the size of this criminogenic effect is small. Table 1 also indicates some notable differences in the Mz when effect sizes are broken down into different subgroups. As should be expected, the overall Mz for effect sizes drawn from bivariate models ($Mz = .089, p < .001$) were approximately 35% larger than those

⁷ All subsequent models (i.e., the moderator analyses reported in tables 2 through 4) also include as covariates the grand mean centered versions of sample size and number of effects per study.

from multivariate models ($Mz = .066, p < .001$). Likewise, the mean effect size for those derived from odds-ratios ($Mz = .093, p < .001$) were approximately 52% larger than those drawn from other test statistics (e.g., t - and z -ratios; $Mz = .061, p < .001$).

[Insert Table 2 about here]

Stability of Effects

To assess the impact of sample and methodological variations on effect size estimates, we also conducted a series of moderator analyses, which are presented in Tables 2 through 4. For these analyses, each moderator was assessed separately; that is, only the focal moderator was entered into the equation (along with covariates for sample size and number of effects per study) and analyses were conducted on the entire sample of 959 effect size estimates. The intercepts in these models can be interpreted as the mean effect size estimate when the focal moderator is set at its reference category. Overall, these results illustrate that the criminogenic effect of custodial sanctions is moderately robust across variations in the characteristics of the samples studied, methodologies employed, and specifications of statistical models. There are, however, several important findings to point out.

[Insert Table 3 about here]

Turning first to the analyses of study characteristics (see Table 2), two variables—publication decade and publication type—had single categories that significantly moderated effect size estimates. However, neither is particularly noteworthy given the small number of effect sizes covered by either. In general, then, the criminogenic effects of custodial sanctions on recidivism are robust across the types of funding received, decade of publication, the countries from which data were drawn, and the type of publication results were reported in. Table 3 displays the results of moderator analyses examining the impact of variations in sample and

sentence characteristics on effect size estimates. Several findings are noteworthy. First, looking at the age composition variable, studies examining exclusively juvenile samples reported larger criminogenic effects of custodial sanctions than in studies examining wholly adult samples ($b = .043, p < .01$). Second, the specific types of both non-custodial and custodial sanctions had significant moderating influences on effect sizes. In terms of the various types of non-custodial sanctions examined, custodial sanctions were less criminogenic relative to non-custodial sanctions for those who received intensive probation ($b = -.097, p < .001$), a suspension or dismissal ($b = -.053, p < .01$), or some other non-custodial sanction ($b = -.063, p < .01$) when compared to those receiving probation. Looking at custodial sanction-types, relative to studies examining prison, being sentenced to jail ($b = .074, p < .001$) or juvenile detention ($b = .094, p < .001$) was associated with a larger criminogenic effect of incarceration. Third, variations in the amount of time spent in custody generally had no moderating influence on effect size estimates. The sole exception was that custodial sanctions appeared more criminogenic than non-custodial sanctions for those who spent between six months and less than one year in custody when compared to those who spent one month to less than six months in custody ($b = .062, p < .01$).

[Insert Table 4 about here]

In Table 4, findings related to the possible moderating effects of variations in research design and model specification are presented. In terms of the type of recidivism assessed, effect size estimates were lower for both convictions ($b = -.022, p < .05$) and reincarcerations ($b = -.026, p < .01$) when compared to models examining arrests/charges. The length of post-sanction follow-up, which varied quite widely across the entire sample, did not emerge as a moderating factor in estimates of the effect of custodial sanctions on recidivism. Finally, we also examined how controlling for or matching members of custodial and non-custodial groups on a host of

theoretically relevant confounds impacted effect size estimates. As Table 4 indicates, models which included controls for or matched on age ($b = -.046, p < .001$), gender ($b = -.030, p < .05$), and sentence length ($b = -.067, p < .01$) produced smaller effect sizes than models which did not, while those which controlled/matched on participants' history of substance abuse produced larger effect size estimates ($b = .051, p < .05$).

The results of the bivariate moderator analyses reported thus far suggest that the weak criminogenic effect of custodial sanctions is fairly robust to variation in sample and methodological characteristics. As Lipsey (2003) has noted, however, moderator variables may be correlated with each other and without accounting for this, their observed impacts on effect sizes may be spurious. For example, although it seems likely that accounting for the age of participants would reduce mean effect size estimates, it is also possible that this observation reflects reduced effect sizes observed in multivariate models compared to bivariate models. Likewise, it is surprising that the inclusion of prior record had no impact on effect size estimates (see Table 4). However, this null finding may be due to greater variability in the full sample of effect sizes, as compared to an analyses of multivariate effect sizes only. In order to address this potential moderator confounding, the next step in our analyses was to examine moderators in a multivariate regression model.

[Insert Table 5 about here]

The results from this meta-regression are reported in Table 5. As an initial step in these analyses, we examined correlations between moderators in the full sample of effect size estimates ($N = 959$) to probe for potential issues with multicollinearity. Correlations between several of the statistical control/matching variable indicators were quite strong in the full sample. For example, the bivariate associations between models having included age and race ($r = .74, p$

< .01), prior record ($r = .75, p < .001$), and current offense ($r = .74, p < .001$) indicated the potential for collinearity issues (Tabachnick & Fidell, 2007). Further, these statistical control/matching indicators had strong associations with the dichotomous indicator for bivariate versus multivariate models. For example, the correlation between the statistical model being multivariate and having included age ($r = .82, p < .001$), race ($r = .79, p < .001$), and prior record ($r = .90, p < .001$) suggested problematic collinearity. Variance inflation factors (VIFs) for these variables were also well above the conservative threshold of 4.0, ranging from 1.13 to 7.65 (Fox, 1991). Accordingly, the full-sample meta-regression (Model 1) was estimated without the inclusion of the indicators for statistical controls/matching variables. As Table 5 shows, the results from this multivariate model on the full sample are quite similar to the bivariate moderator analyses reported in Tables 2 through 4. The one substantive difference is that in the bivariate analyses, samples comprised entirely of juveniles tended to have larger criminogenic effects for custodial sanctions when compared to adult samples ($b = .040, p < .01$), whereas in the multivariate context, this effect was no longer significant ($b = .009, p > .05$).

In the second meta-regression model (Table 5, Model 2), we restricted our analyses to the multivariate sample ($N = 431$). In this reduced sample, correlations between the statistical control/matching variable indicators were no longer problematic (e.g., correlation between age and prior record was $-.040, p > .05$) and none of the VIFs exceeded 4.0; thus, these indicators were included in analyses of the multivariate effect size estimates. As the findings in Table 5 suggest, there are several important differences when examining moderators of effect size within the multivariate-only sample. First, the specific types of sanctions examined—both custodial and non-custodial—made less of an impact on effect size estimates in multivariate models. For instance, in contrast to the analyses of the full sample that included bivariate effect size

estimates, multivariate models that examined jail ($b = -.009, p > .05$) and juvenile detention ($b = .027, p > .05$) did not have significantly different effect sizes from those examining prison as a custodial sanction. Second, the type of reoffending measure used by researchers had no significant influence in multivariate models, with effect sizes from studies examining convictions ($b = -.024, p > .05$), reincarceration ($b = -.011, p > .05$), or technical violations ($b = -.070, p > .05$) not differing significantly from those capturing rearrests. Finally, multivariate models which included controls for/matched on gender ($b = .088, p < .01$), pre-incarceration employment ($b = .064, p < .05$), conviction offense ($b = .044, p < .05$), and prior record ($b = .051, p < .05$) tended to have larger effect sizes than models which did not, while accounting for age ($b = -.110, p < .001$) and marital status ($b = -.051, p < .05$) was associated with smaller effect size estimates.

As a final step in our analyses, we sought to identify the mean effect size estimate in studies that met the methodologically “best-case” criteria. As discussed previously, Nagin et al. (2009; see also Villettaz et al., 2015) note that although RCTs are the gold standard for evaluating treatment effectiveness, they are often not viable nor ethical when sanctioning offenders. At a minimum, then, they urge researchers to implement quasi-experimental designs that can account for the influences of offenders’ age, race/ethnicity, gender, current offense, and prior record on both sanction assignment and reoffending. We therefore conducted a supplementary analysis that compared mean effect size estimates drawn from (a) studies that did implement a RCT, or used a quasi-experimental design (e.g., PSM, natural experiment) that accounted for the five factors noted above, against (b) studies that did meet these criteria. As expected, the mean effect size in studies which did not meet these criteria ($Mz = .080, SE = .009, p < .001$) was higher than those which fulfilled the best-case criteria ($Mz = .052, SE = .022, p < .05$) set forth by Nagin et al. (2009). However, these results illustrate that even in studies which met fairly rigorous

methodological criteria, a weak but statistically significant criminogenic effect of custodial sanctions is observed.

Discussion

Although rates of imprisonment in the United States have declined over the last decade (Bureau of Justice Statistics, 2018), the drive to punish offenders via incarceration largely maintains its grip on correctional policy in America (Clear & Frost, 2014; Pfaff, 2017). If one of the goals of incarceration is to deter individuals from reoffending upon release, it is imperative that we understand whether this is indeed the case for the 600,000 inmates who return to their communities each year (Bureau of Justice Statistics, 2019). Individual studies that compare the reoffending of incarcerated offenders to that of others sentenced to alternative, non-custodial sanctions can shed some light on this question. However, the results from such studies often vary substantially from one another due to differences in the types of sanctions examined, characteristics of the offenders sampled, and the rigor of the methodologies employed to compare groups. The current study sought to make sense of the discrepancies in this body of work through the technique of meta-analysis. Although a limited number of such analyses have been conducted previously (Jonson, 2010; Smith et al., 2002; Villettaz et al., 2015), ours improved on this work by (a) conducting the first inclusive, comprehensive review of the extant research in over a decade, including 54 studies released since 2010; (b) drawing on advancements in modeling techniques that could account for the precision and statistical dependence of effect sizes; and (c) examining whether a wide variety of methodological characteristics moderate mean effect sizes.

The key takeaway from these analyses is that being sentenced to a custodial sanction as opposed to a non-custodial alternative has a weak but statistically significant criminogenic effect

on reoffending ($Mz = .080$). These effects were quite robust to methodological differences between studies and individual statistical models nested within studies. Especially when considered in the context of a multivariate meta-regression, characteristics such as the type of publication (e.g., peer-reviewed versus unpublished); the age and gender composition of samples; the types of custodial and non-custodial sanctions examined; and the type of reoffending measured had little to no effect on effect size estimates. This is not to say that effect sizes were wholly invariant; however, the iatrogenic effects of incarceration were observed even when constraining the sample to RCTs and multivariate effect sizes from quasi-experimental studies that accounted for age, race, gender, prior record, and the type of current offense ($Mz = .052$).

Implications for Policy and Future Research

The finding that imprisoning offenders generally leads to worse outcomes than placing them on probation or other non-custodial sanctions raises serious doubts about deterrence being a legitimate *raison d'être* of incarceration. To be sure, there is certainly a subset of high-risk offenders for whom incapacitation is warranted. There are also issues of the offender's prior record, moral culpability, and the will of the community to mete out punishment for violating the social contract that must be contemplated when it comes to correctional policymaking (Bentham, 1948; Feinberg, 1965). However, for groups of offenders who are statistically alike in terms of their demographic, socioeconomic, criminal history, and offense-related characteristics, our results suggest that imprisonment is not an effective deterrent to future criminal involvement (see, e.g., Aizer & Doyle, 2015; Gaes et al., 2016; Mears & Cochran, 2018). Beyond the inability to reduce reoffending, there are other costs borne from incarcerating those who could otherwise serve their sentences in the community. Fiscally, the ratio of dollars spent per offender sentenced

prison compared to probation is approximately 20 to 1 (i.e., \$78.95 versus \$3.42 per day; PEW Center on the States, 2009). Socially, the constant churning of large swaths of neighborhoods' residents in and out of prison has downstream negative consequences on the strength of social networks and economic vitality of communities, the trust of citizens in the judicial system, and rates of crime (Clear, 2007; Kirk, 2016; Kirk & Papachristos, 2011). And for formerly incarcerated persons themselves, there are myriad other consequences, such as financial hardship (Schwartz-Soicher, Geller, & Garfinkel, 2011), poor health outcomes (Massoglia & Pridemore, 2015), and housing insecurity (Geller & Curtis, 2011).

The findings of the present study also raise several important issues for future research on the effectiveness of correctional sanctions. First, echoing the remarks of Nagin et al., (2009), the results of our meta-regression with the sample of multivariate effect sizes (Table 5) indicate that accounting for the age, race, gender, prior record, and current offense of offenders should be the bare minimum in terms of statistical controls/matching criteria. Given the strong associations between such characteristics and both sentencing outcomes and reoffending, statistical estimates that do not account for them are extremely likely to be biased.

Second, from our own reading of the extant literature in preparation for conducting these analyses, we were struck by the fact that the overwhelming majority of studies did not report whether and/or what proportion of their samples took part in treatment. Among the few that did report such information, we observed substantial variability; in some studies, treatment was provided to both the custodial and non-custodial groups (Wodahl, Boman, & Garland, 2015), to the non-custodial group only (van Wormer & Campbell, 2016), or to the custodial group only (Hiller, Knight, & Simpson, 2006). It is now well-known that providing treatment in accordance with the principles of effective intervention (e.g., risk, needs, responsivity) yields sizeable

reductions in reoffending (Bonta & Andrews, 2017; Landenberger & Lipsey, 2005). Failure to account for participation in such treatment could again lead to biased comparisons of sanction effects. For example, a tenable yet untested hypothesis is that in studies which report a deterrent effect of incarceration, such findings are driven by a larger proportion of inmates having participated in treatment. In this case, the observed crime-reducing effects of imprisonment would be due to rehabilitation rather than deterrence.

Finally, the mechanisms by which imprisonment may increase reoffending largely remain a “black box.” We noted at the outset of the present article that several hypotheses have been advanced, including that prisons (a) are schools of crime that teach and reinforce criminal behavior; (b) lead to a loss of freedoms, exposure to violence, and other aversive experiences that increase strain, anger, and eventual antisocial behavior; or (c) lead to the solidification of the criminal identity through reflected appraisals (i.e., labeling). While each of these hypotheses certainly has theoretical merit, their empirical grounding is in short supply. The schools of crime (Bayer et al., 2009; Harris et al., 2018; Damm & Gorinas, 2016; Ouss, 2011) and strain (Listwan et al., 2013; Morris et al., 2012; Zweig et al., 2015) hypotheses have each been tested only a handful of times. And while some studies have shown that being labeled as a felon or criminal justice processing in general leads to increased reoffending (e.g., Chiricos, Barrick, Bales, & Bontrager, 2007; Gatti, Tremblay, & Vitaro, 2009), the prediction of labeling theory that the stigma of an arrest or incarceration record leads to identity change and increased reoffending has never been adequately tested. The incarceration of millions of Americans is likely to continue for years to come. It is therefore incumbent upon criminologists and policymakers to understand precisely how incarceration can be made less criminogenic than it currently is. If it turns out, for example, that homophily in cell assignments has a substantive influence on the likelihood and

type of reoffending (e.g., Damm & Gorinas, 2016; Ouss, 2011), knowledge on the logistics and impacts of inmates' assignments to specific types of prisons, units, and cells themselves would need to be expanded significantly.

In closing, there is much yet to be learned about the impacts of incarceration on offenders. The findings of the current study, consistent with the prior meta-analytic reviews of the research, suggest that we can be fairly certain that sentencing marginal offenders to prison when they could otherwise be supervised in the community tends not to produce the intended outcome of deterrence. It is evidently both untenable and unwise to place all offenders on community supervision, but the current study suggests that much could be done to unburden the currently over-encumbered and financially strained correctional system with little consequence to public safety. But for the sake of the public and those offenders who undoubtedly should be imprisoned, we also need to know more about how social learning and violence within institutions impact offending trajectories; whether providing more freedoms and improved conditions buffers against these effects; and whether some types of offenders are more negatively impacted by incarceration than others. Further understanding of these issues would enable us to provide offenders with the proper treatment, services, and support they need to address the factors contributing to their criminal behavior.

References

(* Denotes inclusion in meta-analysis)

- *Aarten, P. G. M., & Denkers, A. (2014). Suspending re-offending? Comparing the effects of suspended prison sentences and short-term imprisonment on recidivism in the Netherlands. *European Journal of Criminology*, 11, 702-722.
- *Aarten, P. G. M., Denkers, A., Borgers, M. J., & van der Laan, P. H. (2015). Reconviction rates after suspended sentences: Comparison of the effects of different types of suspended sentences on reconviction in the Netherlands. *International Journal of Offender Therapy and Comparative Criminology*, 59, 143-158.
- *Abrams, D. S. (2010). *Building criminal capital vs specific deterrence: The effect of incarceration length on recidivism* (Unpublished working paper). Retrieved from <https://www.law.upenn.edu/cf/faculty/dabrams/workingpapers/AbramsRecidivism.pdf>
- *Adkins, G., Huff, D., & Stageberg, P. (2000). *The Iowa sex offender registry and recidivism*. Des Moines, IA: Iowa Department of Human Rights.
- *Aizer, A., & Doyle, J. J. Jr. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *The Quarterly Journal of Economics*, 130, 759-804.
- *Andersen, S. H. (2015). Serving time or serving the community? Exploiting a policy reform to assess the causal effects of community service on income, social benefit dependency and recidivism. *Journal of Quantitative Criminology*, 31, 537-563.
- *Babcock, J. C., & Steiner, R. (1999). The relationship between treatment, incarceration, and recidivism of battering: A program evaluation of Seattle's Coordinated Community Response to Domestic Violence. *Journal of Family Psychology*, 13, 46-59.
- *Babst, D. V., & Mannering, J. W. (1965). Probation versus imprisonment for similar types of offenders: A comparison by subsequent violations. *Journal of Research in Crime and Delinquency*, 2, 60-71.
- *Bales, W. D., & Piquero, A. R. (2012). Assessing the impact of imprisonment on recidivism. *Journal of Experimental Criminology*, 8, 71-101.
- *Bartels, L. (2009). The weight of the sword of Damocles: A reconviction analysis of suspended sentences in Tasmania. *The Australian and New Zealand Journal of Criminology*, 42, 72-100.
- *Barton, W. H., & Butts, J. A. (1990). Viable options: Intensive supervision programs for juvenile delinquents. *Crime & Delinquency*, 36, 238-256.
- Bayer, P., Hjalmarsson, R., & Pozen, D. (2009). Building criminal capital behind bars: Peer effects in juvenile corrections. *The Quarterly Journal of Economics*, 124, 105-147.
- Beccaria, C. (1764/1963). *Essays on crimes and punishments*. New York: Macmillan Publishing Company.

- Becker, B. J. (2000). *Handbook of applied multivariate statistics and mathematical modeling*. San Diego, CA: Academic Press.
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76, 169-217.
- *Bell, I. (2011). *2011 compendium of re-offending statistics and analysis*. United Kingdom: Ministry of Justice.
- *Bell, I. (2012). *2012 compendium of re-offending statistics and analysis*. United Kingdom: Ministry of Justice.
- Bentham, J. (1948). *An Introduction to the principles of morals and legislation*. Oxford: Basil Blackwell.
- *Bergman, G. R. (1976). *The evaluation of an experimental program designed to reduce recidivism among second felony criminal offenders* (Unpublished doctoral dissertation). Wayne State University, Michigan.
- *Bhuller, M., Dahl, G. B., Løken, K. V., & Mogstad, M. (2016). *Incarceration, recidivism and employment* (Unpublished working paper). Cambridge, MA: National Bureau of Economic Research. Retrieved from <https://www.nber.org/papers/w22648.pdf>
- Blalock, H. M. (1972). *Social statistics* (2nd Ed.). New York: McGraw-Hill.
- Blevins, K. R., Johnson Listwan, S., Cullen, F. T., & Lero Jonson, C. (2010). A general strain theory of prison violence and misconduct: An integrated model of inmate behavior. *Journal of Contemporary Criminal Justice*, 26, 148-166.
- Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th Ed.). New York: Routledge.
- *Bonta, J., Wallace-Capretta, S., & Rooney, J. (1998). *Restorative justice: An evaluation of the Restorative Resolutions Project*. Ottawa, ON: Solicitor General of Canada.
- *Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000a). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior*, 27, 312-329.
- *Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000b). Can electronic monitoring make a difference? An evaluation of three Canadian programs. *Crime & Delinquency*, 46, 61-75.
- *Bontrager Ryon, S., Winokur Early, K., & Kosloski, A. E. (2017). Community-based and family-focused alternatives to incarceration: A quasi-experimental evaluation of interventions for delinquent youth. *Journal of Criminal Justice*, 51, 59-66.
- *Bontrager Ryon, S., Winokur Early, K., Hand, G., & Chapman, S. (2013). Juvenile justice interventions: System escalation and effective alternatives to residential placement. *Journal of Offender Rehabilitation*, 52, 358-375.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Sussex, UK: John Wiley & Sons Ltd.

- *Brennan, P. A., & Mednick, S. A. (1994). Learning theory approach to the deterrence of criminal recidivism. *Journal of Abnormal Psychology, 103*, 430-440.
- *Brownlee, I. D. (1995). Intensive probation with young adult offenders. *British Journal of Criminology, 35*, 599-612.
- *Bucklen, K. B. (2014). *The impact of returning technical parole violators to prison: A deterrent, null, or criminogenic effect* (Unpublished doctoral dissertation). College Park, MD: University of Maryland, College Park.
- Bureau of Justice Statistics (2018). *2018 update on prisoner recidivism: A 9-year follow-up period (2005-2014)*. Washington, DC: Bureau of Justice Statistics.
- Bureau of Justice Statistics (2019). *Prisoners in 2017*. Washington, DC: Bureau of Justice Statistics.
- *Burns, J. C., & Vito, G. F. (1995). An impact analysis of the Alabama boot camp program. *Federal Probation, 59*, 63-67.
- *Calhoun, K., Etheridge, V., Flinchum, T., Gallagher, A., Hevener, G., & Katzenelson, S. (2009). *Juvenile recidivism study: FY 2004/05 juvenile sample*. Raleigh, NC: North Carolina Sentencing and Policy Advisory Commission.
- *Calhoun, K., Etheridge, V., Flinchum, T., Gallagher, A., Hevener, G., Lagos, D., & Katzenelson, S. (2010). *Correctional program evaluation: Offenders placed on probation or released from prison in fiscal year 2005/06*. Raleigh, NC: North Carolina Sentencing and Policy Advisory Commission.
- *Calhoun, K., Etheridge, V., Flinchum, T., Hevener, G., Katzenelson, S., & Moore-Guerra, M. (2008). *Correctional program evaluation: Offenders placed on probation or released from prison in fiscal year 2003/04*. Raleigh, NC: North Carolina Sentencing and Policy Advisory Commission.
- *Caudy, M. S., Tillyer, M. S., & Tillyer, R. (2018). Jail versus probation: A gender-specific test of differential effectiveness and moderators of sanction effects. *Criminal Justice and Behavior, 45*, 949-968.
- Chalfin, A., & McCrary, J. (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature, 55*, 5-48.
- Chiricos, T., Barrick, K., Bales, W. D., & Bontrager, S. (2007). The labeling of convicted felons and its consequences for recidivism. *Criminology, 45*, 547-581.
- *Cid, J. (2009). Is imprisonment criminogenic? A comparative study of recidivism rates between prison and suspended prison sanctions. *European Journal of Criminology, 6*, 459-380.
- Clear, T. R. (2007). *Imprisoning communities: How mass incarceration makes disadvantaged neighborhoods worse*. New York: Oxford University Press.
- Clear, T. R., & Frost, N. A. (2014). *The punishment imperative: The rise and failure of mass incarceration in America*. New York: New York University Press.

- *Cochran, J. C., Mears, D. P., & Bales, W. D. (2014). Assessing the effectiveness of correctional sanctions. *Journal of Quantitative Criminology*, 30, 317-347.
- *Cohen, B., Eden, R., & Lazar, A. (1991). The efficacy of probation versus imprisonment in reducing recidivism of serious offenders in Israel. *Journal of Criminal Justice*, 19, 263-270.
- *Connecticut Department of Corrections. (2001). *Recidivism in Connecticut report – Final*. Hartford, CT: Connecticut Department of Corrections.
- *Craddock, A., Flinchum, T., Gallagher, A., Hall, M., Hevener, G., Katzenelson, S., & Perdue, S. (2012). Correctional program evaluation: Offenders placed on probation or released from prison in fiscal year 2008/09. Raleigh, NC: North Carolina Sentencing and Policy Advisory Commission.
- Cullen, F. T., Lero Jonson, C., & Nagin, D. S. (2010). Prisons do not reduce recidivism: The high cost of ignoring science. *The Prison Journal*, 91, 48S-65S.
- Damm, A. P., & Gorinas, C. (2016). *Prison as a criminal school: Peer effects and criminal learning behind bars*. The Rockwool Foundation Research Unit Study Paper No. 105. Copenhagen, Denmark: The Rockwool Foundation Research Unit.
- *DeJong, C. (1997). Survival analysis and specific deterrence: Integrating theoretical and empirical models of recidivism. *Criminology*, 35, 561-575.
- *DeLisi, M., Hochstetler, A., Jones-Johnson, G., Caudill, J. W., & Marquart, J. W. (2011). The road to murder: The enduring criminogenic effects of juvenile confinement among a sample of adult career criminals. *Youth Violence and Juvenile Justice*, 9, 207-221.
- *DeYoung, D. J. (1997). An evaluation of the effectiveness of alcohol treatment, drive license actions and jail terms in reducing drunk driving recidivism in California. *Addiction*, 92, 989-997.
- *Erwin, B. S. (1986). Turning up the heat on probationers in Georgia. *Federal Probation*, 50, 17-24.
- Eysenck, H. J. (1984). Meta-analysis: An abuse of research integration. *Journal of Special Education*, 18, 41-59.
- Feinberg, J. (1965). The expressive function of punishment. *The Monist*, 49, 3-15.
- *Flinchum, T., & Hevener, G. (2011). *Juvenile recidivism study: FY 2006/07 juvenile sample*. Raleigh, NC: North Carolina Sentencing and Policy Advisory Commission.
- *Flinchum, T., Dial, R., King, J., Wesoloski, J., Whitaker, B., & Kirk, S. (2018). *Correctional program evaluation: Offenders placed on probation or released from prison, fiscal year 2015*. Raleigh, NC: North Carolina Sentencing and Policy Advisory Commission.
- *Flinchum, T., Hall, M., Hevener, G., & Wesoloski, J. (2016). *Correctional program evaluation: Offenders placed on probation or released from prison in FY 2013*. Raleigh, NC: North Carolina Sentencing and Policy Advisory Commission.

- *Flinchum, T., Hevener, G., Katzenelson, S., & Wesoloski, J. (2014). *Correctional program evaluation: Offenders placed on probation or released from prison in fiscal year 2010/11*. Raleigh, NC: North Carolina Sentencing and Policy Advisory Commission.
- Fox, J. (1991). *Regression diagnostics*. Newbury Park, CA: SAGE.
- *Freiburger, T. L., & Iannacchione, B. M. (2011). An examination of the effect of imprisonment on recidivism. *Criminal Justice Studies*, 24, 369-379.
- *Gaes, G. G., Bales, W. D., & Scaggs, S. J. A. (2016). The effect of imprisonment on recommitment: An analysis using exact, coarsened exact, and radius matching with the propensity score. *Journal of Experimental Criminology*, 12, 143-158.
- Garland, D. (2001). *The culture of control: Crime and social order in contemporary society*. Chicago, IL: University of Chicago Press.
- Gatti, U., Tremblay, R. E., & Vitaro, F. (2009). Iatrogenic effect of juvenile justice. *The Journal of Child Psychology and Psychiatry*, 50, 991-998.
- *Geerken, M. R., & Hayes, H. D. (1993). Probation and parole: Public risk and the future of incarceration alternatives. *Criminology*, 31, 549-564.
- Geller, A., & Curtis, M. A. (2011). A sort of homecoming: Incarceration and the housing security of urban men. *Social Science Research*, 40, 1196-1213.
- Ghandnoosh, N. (2019). *U.S. prison population trends: Massive buildup and modest decline*. Washington, DC: The Sentencing Project.
- *Gibbs, B. R., Lytle, R., & Wakefield, W. (2019). Outcome effects on recidivism among drug court participants. *Criminal Justice and Behavior*, 46, 115-135.
- *Gilman, A. B., Hill, K. G., & Hawkins, J. D. (2015). When is a youth's debt to society paid? Examining the long-term consequences of juvenile incarceration for adult functioning. *Journal of Developmental and Life Course Criminology*, 1, 33-47.
- *Gordon, M. A., & Glaser, D. (1991). The use and effects of financial penalties in municipal courts. *Criminology*, 29, 651-676.
- *Gottfredson, D. C., & Barton, W. H. (1993). Deinstitutionalization of juvenile offenders. *Criminology*, 31, 591-611.
- *Harding, D. J., Morenoff, J. D., Nguyen, A. P., & Bushway, S. D. (2017). Short- and long-term effects of imprisonment on future felony convictions and prison admissions. *Proceedings of the National Academy of Sciences*, 114, 11103-11108.
- Harris, H. M., Nakamura, K., & Bucklen, K. B. (2018). Do cellmates matter? A causal test of the schools of crime hypothesis with implications for differential association and deterrence theories. *Criminology*, 56, 87-122.
- *Hedderman, C., & Jolliffe, D. (2015). The impact of prison for women on the edge: Paying the price for wrong decisions. *Victims & Offenders*, 10, 152-178.

- *Henneguelle, A., Monnery, B., & Kensey, A. (2016). Better at home than in prison? The effects of electronic monitoring on recidivism in France. *Journal of Law and Economics*, 59, 629-667.
- *Hiller, M. L., Knight, L., & Simpson, D. D. (2006). Recidivism following mandated residential substance abuse treatment for felony probationers. *The Prison Journal*, 86, 230-241.
- *Hjalmarsson, R. (2009). Juvenile jails: A path to the straight and narrow or to hardened criminality? *The Journal of Law and Economics*, 52, 779-809.
- Hox, J. J. (2017). *Multilevel analysis: Techniques and applications* (3rd Ed.). New York: Routledge.
- Hox, J. J., & de Leeuw, E. D. (2003). Multilevel models for meta-analysis. In S. P. Reise & N. Duan (Eds.), *Multilevel modeling: methodological advances, issues, and applications* (pp. 90-111). Mahwah, NJ: Lawrence Erlbaum.
- *Hunt, K. S., & Dumville, T. (2016). *Recidivism among federal offenders: A comprehensive overview*. Washington, DC: United States Sentencing Commission.
- *Hunt, K. S., Iaconetti, M. J., & Maass, K. T. (2019). *Recidivism among federal violent offenders*. Washington, DC: United States Sentencing Commission.
- *Jolliffe, D., & Hedderman, C. (2015). Investigating the impact of custody on reoffending using propensity score matching. *Crime & Delinquency*, 61, 1051-1077.
- *Jones, M., & Ross, D. L. (1997a). Electronic house arrest and boot camp in North Carolina: Comparing recidivism. *Criminal Justice Policy Review*, 8, 383-403.
- *Jones, M., & Ross, D. L. (1997b). Is less better? Boot camp, regular probation and rearrest in North Carolina. *Journal of Criminal Justice*, 21, 147-161.
- *Jones, P. R. (1991). The risk of recidivism: Evaluating the public-safety implications of a community corrections program. *Journal of Criminal Justice*, 19, 49-66.
- Jonson, C. L. (2010). *The impact of imprisonment on offending: A meta-analysis* (Unpublished doctoral dissertation). Cincinnati, OH: University of Cincinnati.
- *Killias, M., Aebi, M., & Ribeaud, D. (2000). Does community service rehabilitate better than short-term imprisonment? Results of a controlled experiment. *The Howard Journal*, 39, 40-57.
- *Killias, M., Gilliéron, G., Villard, F., & Poggia, C. (2010). How damaging is imprisonment in the long-term? A controlled experiment comparing long-term effects of community service and short custodial sentences on re-offending and social integration. *Journal of Experimental Criminology*, 6, 115-130.
- Kirk, D. S. (2016). Prisoner reentry and the reproduction of legal cynicism. *Social Problems*, 63, 222-243.
- Kirk, D. S., & Papachristos, A. V. (2011). Cultural mechanisms and the persistence of neighborhood violence. *American Journal of Sociology*, 116, 1190-1233.

- *Klement, C. (2015). Comparing the effects of community service and imprisonment on reconviction: Results from a quasi-experimental Danish study. *Journal of Experimental Criminology*, 11, 237-261.
- *Kraus, J. (1974). A comparison of corrective effects of probation and detention on male juvenile offenders. *British Journal of Criminology, Delinquency and Deviant Social Behaviour*, 14, 49-62.
- *Kraus, J. (1978). Remand in custody as a deterrent in juvenile jurisdiction. *British Journal of Criminology, Delinquency and Deviant Social Behaviour*, 18, 285-289.
- *Kraus, J. (1981). The effects of committal to a special school for truants. *International Journal of Offender Therapy and Comparative Criminology*, 25, 130-138.
- Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: SAGE.
- Landenberger, N. A., & Lipsey, M. W. (2005). The positive effect of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1, 451-476.
- *Lin, J. (2007). *Exploring the impact of institutional placement on the recidivism of delinquent youth*. National Institute of Justice.
- Listwan, S. J., Sullivan, C. J., Agnew, R., Cullen, F. T., & Colvin, M. (2013). The pains of imprisonment revisited: The impact of strain on inmate recidivism. *Justice Quarterly*, 30, 144-168.
- *Loeffler, C. E. (2013). Does imprisonment alter the life course? Evidence on crime and employment from a natural experiment. *Criminology*, 51, 137-166.
- *MacKenzie, D. L. (1991). The parole performance of offenders released from shock incarceration (boot camp prisons): A survival time analysis. *Journal of Quantitative Criminology*, 7, 213-236.
- *MacKenzie, D. L., & Shaw, J. W. (1993). The impact of shock incarceration on technical violations and new criminal activities. *Justice Quarterly*, 10, 463-487.
- Massoglia, M., & Pridemore, W. A. (2015). Incarceration and health. *Annual Review of Sociology*, 41, 291-310.
- *Mbuba, J. M. (2004). *Juvenile recidivism: An analysis of race and other socio-demographic predictors within three intervention modalities in the state of Louisiana* (Unpublished doctoral dissertation). Louisiana State University.
- *McAlister, S., Officer, K., & Sanchagrin, K. (2019). *Oregon recidivism analysis*. Salem, OR: Oregon Criminal Justice Commission.
- *Mears, D. P., & Cochran, J. C. (2018). Progressively tougher sanctioning and recidivism: Assessing the effects of different types of sanctions. *Journal of Research in Crime and Delinquency*, 55, 194-241.
- *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.

- *Ministry of Justice. (2013). *2013 compendium of re-offending statistics and analysis*. United Kingdom: Ministry of Justice.
- *Mitchell, O., Cochran, J. C., Mears, D. P., & Bales, W. D. (2017). Examining prison effects on recidivism: A regression discontinuity approach. *Justice Quarterly*, 34, 571-596.
- Morris, R. G., Carriaga, M. L., Diamond, B., Leeper Piquero, N., & Piquero, A. R. (2012). Does prison strain lead to prison misbehavior? An application of general strain theory to inmate misconduct. *Journal of Criminal Justice*, 40, 194-201.
- *Mueller-Smith, M. (2014). *The criminal and labor market impacts of incarceration* (Unpublished working paper). Retrieved from <https://sites.lsa.umich.edu/mgms/wp-content/uploads/sites/283/2015/09/incar.pdf>
- *Muiluvuori, M. L. (2001). Recidivism among people sentenced to community service in Finland. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 2, 72-82.
- Myers, W., Turanovic, J. J., Lloyd, K. M., & Pratt, T. C. (2020). The victimization of LGBTQ students at school: A meta-analysis. *Journal of School Violence* (Online First). <https://doi.org/10.1080/15388220.2020.1725530>
- Nagin, D. S., Cullen, F. T., & Lero Jonson, C. (2009). Imprisonment and reoffending. *Crime and Justice*, 38, 115-200.
- National Academy of Sciences (2014). *The growth of incarceration in the United States: Exploring causes and consequences*. Washington, DC: The National Academies Press.
- Nellis, A., & King, R. S. (2009). *No exit: The expanding use of life sentences in America*. Washington, DC: The Sentencing Project.
- *New Mexico Sentencing Commission (2014). *Updated exploratory sex offender recidivism study: 2004-2006 probation sentence and prison release cohorts*. Albuquerque, NM: New Mexico Sentencing Commission.
- *Nieuwbeerta, P., Nagin, D. S., & Blokland, A. A. J. (2009). Assessing the impact of first-time imprisonment on offenders' subsequent criminal career development: A matched samples comparison. *Journal of Quantitative Criminology*, 25, 227-257.
- *Nirel, R., Landau, S. F., Sebba, L., & Sagiv, B. The effectiveness of service work: An analysis of recidivism. *Journal of Quantitative Criminology*, 13, 73-92.
- *Nunes, K. L., Firestone, P., Wexler, A. F., Jensen, T. L., & Bradford, J. M. (2007). Incarceration and recidivism among sexual offenders. *Law and Human Behavior*, 31, 305-318.
- *Office of the Legislative Auditor. (1997). *Recidivism of adult felons: A program evaluation report*. St. Paul, MN: Office of the Legislative Auditor.
- *Oregon Department of Corrections. (2002). *The effectiveness of community-based sanctions in reducing recidivism*. Salem, OR: Oregon Department of Corrections.

- Ouss, A. (2011). *Prison as a school of crime: Evidence from cell-level interactions* (Unpublished working paper). Social Science Research Network. Retrieved from <https://dx.doi.org/10.2139/ssrn.1989803>
- Pager, D. (2007). *Marked: Race, crime, and finding work in an era of mass incarceration*. Chicago, IL: University of Chicago Press.
- Persoskie, A., & Ferrer, R. A. (2017). A most odd ratio: Interpreting and describing odds ratios. *American Journal of Preventive Medicine*, 52, 224-228.
- Pew Center on the States (2009). *One in 31: The long reach of American corrections*. Washington, DC: The Pew Charitable Trusts.
- Pfaff, J. F. (2017). *Locked in: The true causes of mass incarceration and how to achieve real reform*. New York: Basic Books.
- Pratt, T. C., & Cullen, F. T. (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology*, 38, 931-964.
- Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L. H., & Madensen, T. D. (2006). The empirical status of deterrence theory: A meta-analysis. In F. T. Cullen, J. P. Wright, and K. R. Blevins (eds.), *Taking stock: The status of criminological theory - Advances in criminological theory, Vol. 15* (pp. 367-395). Piscataway, NJ: Transaction Publishers.
- Pratt, T. C., Turanovic, J. J., Fox, K. A., & Wright, K. A. (2014). Self-control and victimization: A meta-analysis. *Criminology*, 52, 87-116.
- Pyrooz, D. C., Turanovic, J. J., Decker, S. H., & Wu, J. (2016). Taking stock of the relationship between gang membership and offending: A meta-analysis. *Criminal Justice and Behavior*, 43, 365-397.
- *Robert, L., Maes, E., Blokland, A., & Wermink, H. (2017). 'Virtual' versus 'real' prison: Which is best? Comparing the re-incarceration rates after electronic monitoring and imprisonment in Belgium. In A. Blokland, & V. van der Geest (Eds.), *The Routledge international handbook of life-course criminology* (pp. 417-435). New York: Routledge.
- *Rodriguez-Menés, J., & Creighton, M. (2017). The impact of imprisonment on IPV offenders' risks of recidivism: An application of two natural experiments in the city of Barcelona. In A. Blokland, & V. van der Geest (Eds.), *The Routledge international handbook of life-course criminology* (pp. 397-416). New York: Routledge.
- *Roeger, L. S. (1994). The effectiveness of criminal justice sanctions for Aboriginal offenders. *The Australian and New Zealand Journal of Criminology*, 27, 264-281.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231-244). New York: Russell Sage Foundation.
- Sampson, R. J., & Laub, J. H. (1993). *Crime in the making: Pathways and turning points through life*. Cambridge, MA: Harvard University Press.
- *Scarpitti, F. R., & Stephenson, R. M. (1968). A study of probation effectiveness. *The Journal of Criminal Law, Criminology, and Police Science*, 59, 361-369.

- *Schneider, A. L. (1986). Restitution and recidivism rates of juvenile offenders: Results from four experimental studies. *Criminology*, 24, 533-552.
- Schwartz-Soicher, O., Geller, A., & Garfinkel, I. (2011). The effect of paternal incarceration on material hardship. *Social Service Review*, 85, 447-473.
- *Schweitzer, M., Labrecque, R. M., & Smith, P. (2017). Reinvesting in the lives of youth: A targeted approach to reducing recidivism. *Criminal Justice Policy Review*, 28, 207-219.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Belmont, CA: Wadsworth.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review*, 17, 881-901.
- *Sheldon, R. G. (1997). *An assessment of the Detention Diversion Advocacy Project: Final report*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.
- *Sirén, R., & Savolainen, J. (2013). No evidence of specific deterrence under penal moderation: Imprisonment and recidivism in Finland. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 14, 80-97.
- *Smith, L. G., & Akers, R. L. (1993). A comparison of recidivism of Florida's community control and prison: A five-year survival analysis. *Journal of Research in Crime and Delinquency*, 30, 267-292.
- Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore, MD: Johns Hopkins.
- Smith, P., Goggin, C., & Gendreau, P. (2002). *The effects of prison sentences and intermediate sanctions on recidivism: General effects and individual differences*. Ottawa, ON: Solicitor General of Canada.
- *Spohn, C. (2007). The deterrent effect of imprisonment and offenders' stakes in conformity. *Criminal Justice Policy Review*, 18, 31-50.
- *Spohn, C., & Holleran, D. (2002). The effect of imprisonment on recidivism rates of felony offenders: A focus on drug offenders. *Criminology*, 40, 329-357.
- *Steiner, B., & Giacomazzi, A. L. (2007). Juvenile waiver, boot camp, and recidivism in a Northwestern state. *The Prison Journal*, 87, 227-240.
- *Sweeten, G., & Apel, R. (2007). *Incarceration and the transition to adulthood*. National Poverty Center.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th Ed.). Boston, MA: Allyn and Bacon.
- *Texas Legislative Board (2019). *Statewide criminal justice and juvenile justice recidivism and revocation rates*. Austin, TX: Texas Legislative Budget Board.
- *Trevena, J., & Poynton, S. (2016). Does a prison sentence affect future domestic violence reoffending? *Contemporary Issues in Crime and Justice*, 190, 1-12.

- Turanovic, J. J., & Pratt, T. C. (2020). Meta-analysis in criminology and criminal justice: Challenging the paradigm and charting a new path forward. *Justice Evaluation Journal* (Online First). <https://doi.org/10.1080/24751979.2020.1775107>
- *Ulmer, J. T. (2001). Intermediate sanctions: A comparative analysis of the probability and severity of recidivism. *Sociological Inquiry*, 71, 164-193.
- *Van Der Werff, C. (1981). Recidivism and special deterrence. *British Journal of Criminology, Delinquency and Deviant Social Behaviour*, 21, 136-147.
- *Van Ness, S. R. (1992). Intensive probation versus prison outcomes in Indiana: Who could benefit? *Journal of Contemporary Criminal Justice*, 8, 351-364.
- *van Wormer, J. G., & Campbell, C. (2016). Developing an alternative juvenile programming effort to reduce detention overreliance. *OJJDP Journal of Juvenile Justice*, 5(2), 12-30.
- Vartanian, L. R., Schwartz, M. B., & Bronwell, K. D. (2007). Effects of soft drink consumption on nutrition and health: A systematic review and meta-analysis. *American Journal of Public Health*, 97, 667-675.
- Vera Institute (2017). *The price of prisons: Examining state spending trends, 2010-2015*. Brooklyn, NJ: Vera Institute of Justice.
- *Villanueva, L., & Cuervo, K. (2018). The impact of juvenile educational measures, confinement centers, and probation on adult recidivism. *International Journal of Offender Therapy and Comparative Criminology*, 62, 4108-4123.
- Villetaz, P., Gillieron, G., & Killias, M. (2015). *The effects on re-offending of custodial vs. non-custodial sanctions: An updated systematic review of the state of knowledge*. Oslo, Norway: The Campbell Collaboration.
- Villetaz, P., Killias, M., & Zoder, I. (2006). *The effects of custodial vs. non-custodial sentences on re-offending: A systematic review of the state of knowledge*. Oslo, Norway: The Campbell Collaboration.
- *Vito, G. F., & Allen, H. E. (1981). Shock probation in Ohio: A comparison of outcomes. *International Journal of Offender Therapy and Comparative Criminology*, 25, 70-75.
- *Wang, J. J., & Poynton, S. (2017). Intensive correction orders versus short prison sentence: A comparison of re-offending. *Contemporary Issues in Criminal Justice*, 207, 1-19.
- *Weatherburn, D., Vignaendra, S., & McGrath, A. (2009). The specific deterrent effect of custodial penalties on juvenile re-offending. *Contemporary Issues in Crime and Justice*, 132, 1-8.
- Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., ... & Fordwood, S. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist*, 72, 79-117.

- *Wermink, H., Blokland, A., Nieuwbeerta, P., Nagin, D. S., & Tollenaar, N. (2010). Comparing the effects of community service and short-term imprisonment on recidivism: A matched samples approach. *Journal of Experimental Criminology*, 6, 325-349.
- *Wheeler, G. R., & Hissong, R. V. (1988). A survival time analysis of criminal sanctions for misdemeanor offenders: A case for alternatives to incarceration. *Evaluation Review*, 12, 510-527.
- *Wheeler, G. R., & Hissong, R. V. (1988). Effects of criminal sanctions on drunk drivers: Beyond incarceration. *Crime & Delinquency*, 34, 29-42.
- *Wiebush, R. G. (1993). Juvenile intensive supervision: The impact on felony offenders diverted from institutional placement. *Crime & Delinquency*, 39, 68-89.
- *Wodahl, E. J., Boman, J. H. IV., & Garland, B. E. (2015). Responding to probation and parole violations: Are jail sanctions more effective than community-based graduated sanctions? *Journal of Criminal Justice*, 43, 242-250.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Newbury Park, CA: SAGE.
- *Wong, T. (2017). *State of Hawaii, FY 2013 cohort: 2016 recidivism update*. Honolulu, HI: Hawaii State Department of Health.
- *Wong, T. (2018). *State of Hawaii, FY 2014 cohort: 2017 recidivism update*. Honolulu, HI: Hawaii State Department of Health.
- *Wooldredge, J. (1988). Differentiating the effects of juvenile court sentences on eliminating recidivism. *Journal of Research in Crime and Delinquency*, 25, 264-300.
- *Wooldredge, J. (2007). Convicting and incarcerating felony offenders of intimate assault and the odds of new assault charges. *Journal of Criminal Justice*, 35, 379-389.
- World Prison Brief (2018). *World prison population list* (12th Ed.). London, UK: Institute for Criminal Policy Research.
- *Wright, D. T., & Mays, G. L. (1998). Correctional boot camps, attitudes, and recidivism. *Journal of Offender Rehabilitation*, 28, 71-87.
- Zweig, J. M., Yahner, J., Visser, C. A., & Lattimore, P. K. (2015). Using general strain theory to explore the effects of prison victimization experiences on later offending and substance use. *The Prison Journal*, 95, 84-113.

Table 1. Mean Estimates for the Effect of Custodial Sanctions on Recidivism

Model Estimation	Mean Effect Size Estimate	95 Percent CI
Overall effect size (959)	.080***	.062 to .097
Bivariate effect size (528)	.089***	.063 to .116
Multivariate effect size (431)	.066***	.046 to .086
OR-derived effect size (607)	.093***	.067 to .119
Non-OR-derived effect size (352)	.061***	.043 to .079

Notes: Number of effect sizes are in parentheses.

Abbreviations: OR = odds-ratio; CI = confidence interval.

*** $p < .001$ (two-tailed test).

Table 2. Bivariate Moderator Analyses for the Impact of Study Characteristics on Effect Size Estimates

Moderator Variable	Coefficient	SE	z-Value	Intercept
Funding source				.082***
Not reported (329)				
State/provincial funded (361)	.013	.024	.55	
Federally funded (180)	-.008	.022	-.36	
Privately funded (17)	-.040	.047	-.85	
Unfunded (34)	-.035	.035	-1.00	
Other funding source (34)	.018	.069	.26	
Publication decade				.085***
2010s (636)				
2000s (127)	-.019	.023	-.82	
1990s (144)	-.029	.022	-1.29	
1980s (28)	.028	.036	.77	
1970s (5)	.160	.069	2.32*	
1960s (19)	.014	.061	.23	
Study location				.068***
United States (751)				
Canada (12)	.019	.051	.37	
United Kingdom (51)	.032	.038	.86	
Australia (33)	.048	.038	1.27	
Nordic countries (44)	.003	.038	.08	
The Netherlands (36)	.038	.040	.97	
Other (32)	.055	.033	1.65	
Publication type				.065***
Peer-reviewed article (479)				
State/local report (334)	.053	.024	2.18*	
Federal report (57)	.053	.035	1.50	
Thesis/dissertation (16)	-.038	.047	-.80	
Book chapter (10)	.157	.059	2.65**	
Other (63)	.034	.037	.91	

Notes: Estimates are based on the full sample ($N = 959$). The frequencies of effect size estimates are in parentheses. Models were estimated separately for each moderator.

Abbreviations: SE = standard error.

* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed test).

Table 3. Bivariate Moderator Analyses for the Impact of Sample Characteristics on Effect Size Estimates

Moderator Variable	Coefficient	SE	z-Value	Intercept
Age composition				.067***
Exclusively adults (630)				
Exclusively juveniles (154)	.040	.014	2.89**	
>80% adults (26)	.023	.037	.63	
>80% juveniles (5)	.025	.074	.34	
Missing (136)	.048	.041	1.17	
Gender composition				.066***
Exclusively males (152)				
Exclusively females (30)	.026	.025	1.04	
>80% males (321)	.016	.019	.83	
Mixed (128)	.080	.013	6.38***	
Missing (328)	-.011	.020	-.56	
Non-custodial sanction type				.115***
Probation (518)				
Intensive probation (146)	-.101	.012	-8.42***	
EM/house arrest (30)	-.049	.029	-1.34	
Community service/fine (72)	-.043	.023	-1.92	
Tx-focused NCS (16)	-.023	.039	-.59	
Suspended/dismissal (57)	-.054	.019	-2.85**	
Other (120)	-.062	.020	-3.19*	
Custodial sanction type				.056***
Prison (611)				
Jail (83)	.074	.015	4.82***	
Juvenile detention (104)	.092	.018	5.06***	
Boot camp/shock (36)	-.037	.037	-1.02	
Residential treatment (9)	.018	.062	.30	
Other (116)	.036	.026	1.43	
Time in custody				.076***
1 to <6 months (130)				
<1 month (25)	-.051	.033	-1.53	
6 months to <1 year (54)	.067	.023	2.87**	
1 to <2 years (9)	-.039	.054	-.73	
>= 2 years (26)	-.050	.074	-.69	

Notes: Estimates are based on the full sample ($N = 959$). The frequencies of effect size estimates are in parentheses. Models were estimated separately for each moderator.

Abbreviations: EM = electronic monitoring; NCS = non-custodial sanction; SE = standard error; Tx = treatment.

** $p < .01$; *** $p < .001$ (two-tailed test).

Table 4. Bivariate Moderator Analyses for the Impacts of Research Design and Model Specification on Effect Size Estimates

Moderator Variable	Coefficient	SE	z-Value	Intercept
Recidivism measure				.091***
Arrest/charge (289)				
Conviction (409)	-.022	.009	-2.42*	
Reincarceration (217)	-.025	.009	-2.70**	
Technical violation (15)	.047	.031	1.51	
Other (26)	-.003	.025	-.14	
Length of follow-up				.071***
≤ 1 year (241)				
> 1 year to ≤ 2 years (205)	.011	.010	1.12	
> 2 years to ≤ 3 years (341)	.009	.010	.93	
> 3 years to ≤ 4 years (41)	.025	.027	.92	
> 4 years (125)	.016	.017	.95	
Statistical controls/matching variables				
Age (416)	-.047	.012	-3.93***	.103***
Gender (391)	-.031	.013	-2.47*	.093***
Marital status (64)	-.011	.023	-.49	.081***
Employment status (76)	-.010	.021	-.47	.081***
Education (140)	.016	.015	.99	.077***
Socioeconomic status (65)	.033	.020	1.61	.078***
Race/ethnicity (349)	-.000	.013	-.00	.080***
Conviction offense (310)	-.020	.014	-1.48	.087***
Age at first offense (51)	.019	.022	.83	.077***
Prior record (387)	.001	.012	.11	.079***
Substance abuse (77)	.050	.021	2.41*	.076***
Mental health (36)	-.025	.058	-.43	.080***
Risk level (67)	-.039	.022	-1.80	.085***
Sentence length (95)	-.069	.019	-3.58***	.085***

Notes: Estimates are based on the full sample ($N = 959$). The frequencies of effect size estimates are in parentheses. Models were estimated separately for each moderator.

Abbreviation: SE = standard error.

* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed test).

Table 5. Multivariate Meta-Regression of Effect Size Estimates on Methodological Variations

Moderator Variables	Model 1		Model 2	
	Full Sample		Multivariate Sample	
	Coefficient	SE	Coefficient	SE
Fixed Effects				
Model Intercept	.098***	.038	.098**	.038
Publication type				
Peer-reviewed article				
State/local report	.038	.031	---	---
Federal report	.077	.040	-.038	.036
Thesis/dissertation	-.015	.054	-.072	.038
Book chapter	.187***	.067	.176***	.032
Other	.084	.044	-.002	.040
Publication decade				
2010s				
2000s	-.015	.026	.003	.023
1990s	-.034	.028	-.003	.028
1980s	-.030	.048	-.159**	.059
1970s	.103	.075	---	---
1960s	-.040	.066	---	---
Age composition				
Exclusively adults				
Exclusively juveniles	.009	.023	-.002	.017
>80% adults	.061	.039	.022	.040
>80% juveniles	-.009	.077	-.007	.056
Missing	.056	.048	---	---
Gender composition				
Exclusively males				
Exclusively females	.008	.020	.016	.014
>80% males	-.005	.020	-.031	.022
Mixed	-.067*	.031	-.061*	.027
Missing	-.051	.027	-.030	.024
Non-custodial sanction type				
Probation				
Intensive probation	-.101***	.011	.005	.014
EM/house arrest	-.056	.029	-.045	.027
Community service/fine	-.033	.022	-.003	.022
Tx-focused NCS	-.029	.041	.002	.030
Suspended/dismissal	-.050**	.018	-.047***	.013
Other	-.094***	.022	-.037	.026
Custodial sanction type				
Prison				
Jail	.083***	.014	-.009	.015
Juvenile detention	.094**	.028	.027	.043
Boot camp/shock	-.023	.037	-.154***	.035

Residential treatment	.032	.063	.026	.022
Other	.062*	.028	-.015	.045
Recidivism measure				
Arrest/charge				
Conviction	-.022*	.008	-.024	.014
Reincarceration	-.029**	.009	-.011	.015
Technical violation	.059*	.029	-.070	.041
Other	.005	.023	.006	.017
Statistical controls/matching variables				
Age			-.110***	.026
Gender			.088**	.027
Marital status			-.051*	.024
Employment status			.064*	.029
Education			.008	.028
Socioeconomic status			-.033	.026
Race/ethnicity			-.014	.021
Conviction offense			.044*	.019
Age at first offense			.015	.019
Prior record			.051*	.021
Substance abuse			-.003	.025
Mental health			-.068	.039
Risk level			-.030	.027
Sentence length			-.024	.015
Multivariate statistical model	-.016	.014		
Sample size	.000	.001	.000	.001
Number of effect sizes (Level 2)	.000	.000	.000	.000
Random Effects				
Level 1 (Effect size estimates)				
Variance between models	.00629		.00130	.00059
Variance explained	.000		.759	
Level 2 (Study)				
Variance between studies	.00494		.00174	.00019
Variance explained	.000		.148	

Note: Full sample $N = 959$; Multivariate sample $N = 431$.

Abbreviation: SE = standard error.

* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed test).