Journal of Experimental Criminology
https://doi.org/10.1007/s11292-018-9341-7

A quasi-experimental test of the effects of criminal justice involvement on later mental health

Jessica M. Craig[1] · Alex R. Piquero[2] · Joseph Murray[3] · David P. Farrington[4]

Published online: 04 October 2018
© Springer Nature B.V. 2018

Abstract

Objectives While many criminological theories posit causal hypotheses, many studies fail to use methods that adequately address the three criteria of causality. This is particularly important when assessing the impact of criminal justice involvement on later outcomes. Due to practical and ethical concerns, it is challenging to randomize criminal sanctions, so quasi-experimental methods such as propensity score matching are often used to approximate a randomized design. Based on longitudinal data from the Cambridge Study in Delinquent Development, the current study used propensity score matching to investigate the extent to which convictions and/or incarcerations in the first two decades of life were related to adverse mental health during middle adulthood.
Methods Propensity scores were utilized to match those with and without criminal justice involvement on a wide range of risk factors for offending.
Results The results indicated that there were no significant differences in mental health between those involved in the criminal justice system and those without such involvement.
Conclusions The results did not detect a relationship between justice system involvement and later mental health suggesting that the consequences of criminal justice involvement may only be limited to certain domains.

Keywords Effects of conviction · Effects of incarceration · Labeling theory · Mental health · Propensity score matching

✉ Jessica M. Craig
jessica.craig@unt.edu

[1] University of North Texas, Denton, TX, USA

[2] University of Texas at Dallas, Richardson, TX, USA

[3] Federal University of Pelotas, Pelotas, Brazil

[4] University of Cambridge, Cambridge, UK

Introduction

Many criminological theories hypothesize causal relationships between their key variables of interest and offending. For instance, deterrence theory anticipates that sanctions will decrease future criminal activity by increasing the certainty, celerity, and severity of punishment. Labeling theory, on the other hand, makes a different prediction: that sanctions will backfire and lead to increased criminal activity through their effects on an individual's identity and available opportunities. These are two opposing causal hypotheses regarding the impact of sanctions on later offending; one expects that criminal justice sanctions will cause a decrease in offending while the other claims that the same sanctions will cause an increase in offending. Despite these contradictory claims, both theories have some level of empirical support in the literature (for labeling theory, see Becker 1963; Farrington and Murray 2014; Lemert 1967; for deterrence theory, see Durlauf and Nagin 2010; Nagin 2013).

Perhaps one of the reasons both theories continue to receive interest from criminologists is a function of the empirical challenges associated with answering causal questions. From a methodological standpoint, a randomized experiment would be the ideal choice as it is capable of eliminating more threats to internal validity than other study designs (Shadish et al. 2002). As discussed by Weisburd and Hinkle (2012), non-experimental methods utilize "knowledge" solutions for the issue of confounding variables. In other words, if a researcher is aware of all confounding variables, then s/he will control for those statistically, such as in a regression model. True experimental designs, on the other hand, rely upon randomization of participants to the experimental and control groups to control for confounding variables. The most important feature of random assignment is that the probability of two adequately sized experimental and control groups being systematically different is very low; thus, researchers do not need to identify all potential confounding variables. However, as noted by Apel and Sweeten (2010), there are important practical and ethical constraints faced by researchers in the implementation of these designs. Within the context of the debate between deterrence and labeling theories, it is difficult to randomly assign some study participants to receive criminal justice sanctions while others do not (for an exception, see Klein 1986).

In situations like these, the second best option is quasi-experimentation, or designs that seek to closely replicate random assignment. One such approach is propensity score matching (PSM), which is a knowledge-based solution (Weisburd and Hinkle 2012) that gives researchers the ability to match those who received the "treatment" to those who did not "on a very large number of measured characteristics, including pretreatment outcomes" (Apel and Sweeten 2010: 544). As will be discussed further, this is the method the current study uses to investigate the effects of criminal justice system involvement on later life outcomes among the men in the Cambridge Study in Delinquent Development. Specifically, we focus on the extent to which convictions and/or incarcerations in the first two decades of life are related to adverse mental health functioning during the 30s in a sample of London males. The PSM approach attends to potential selection problems regarding the likelihood of convictions and incarcerations and allows us to address the issue of differential selection at the outset of the relationship between conviction/incarceration and mental health.

Theory and current literature

Prior to discussing the results of the PSM analyses, we first offer a brief review of the literature detailing the effects of criminal sanctions not just on future criminal behavior but on other outcomes such as employment and physical health. Next, we turn to a discussion of the connection between criminal justice system involvement and mental health. Finally, we close with an explanation of propensity score matching, focusing on how it offers a stronger test of the impact of criminal justice sanctions on later outcomes compared with other non-experimental methods.

The theoretical and empirical impact of criminal sanctions

The roots of labeling theory can be traced back to both conflict theory and symbolic interactionism. The conflict perspective focuses on who is labeled and who gets to decide who and what is labeled (Becker 1963). Symbolic interaction arguments, such as those put forth by Garfinkel (1956), Lemert (1967), and Matsueda (1992), center on the impact of perceptions and appraisals on a person's later behavior. Lemert's concept of secondary deviance shifts the focus from what are the potential causes of crime (especially the initial criminal act) to what happens after a criminal sanction is administered. Specifically, once an individual is convicted in court and receives the label of "convicted criminal," there are not only structural opportunities that persons are cut off from, such as certain student loans or careers, but other people's appraisals of the person start to shift, casting the individual's character in a negative light. Over time, individuals come to be seen as "deviant" or "criminal," and may be turned away from their prior prosocial groups only to be accepted by other deviant groups. The label of "criminal" may become a part of the person's identity. Through these changes in opportunities, appraisals, and identities, the labeled person becomes at greater risk for offending and other antisocial behavior (Lemert 1967).

A central tenet of labeling theory, then, is that experience with the criminal justice system—ranging from a mild police contact to an arrest, a trial, conviction, and any potential subsequent punishments—may do more harm than good if it results in a change in identity or available opportunities. Theoretical and empirical research on the effects and experiences associated with justice system experience and involvement has been extensive (see review in Farrington and Murray 2014). Briefly, these adverse justice-related outcomes include, for example, a change in one's self-appraisal as a consequence of the perceived appraisal of others (Matsueda 1992), being blocked from prosocial opportunities as a result of a criminal record (Pager 2003), and especially secondary or continued deviance and offending (see Liberman et al. 2014; Morris and Piquero 2013). In sum, many studies have found some level of support for the adverse impact of a criminal justice sanction (Bernburg and Krohn 2003; Davies and Tanner 2003; Farrington 1977; Farrington et al. 1978; Huizinga and Henry 2008; Klein 1986; Li 1999; Murray et al. 2014; Nagin and Waldfogel 1995; Sweeten 2006). However, this is not true for all studies as some, including those drawing upon stronger methodological designs, have reported null or deterrent effects (Bhati and Piquero 2007; Green and Winik 2010; Loeffler 2013; Loughran et al. 2009; Smith and Gartin 1989).

Several non-experimental designs have suggested that those who were involved in the juvenile justice or criminal justice system also experienced a decrease in their later

educational attainment and employment (Bernburg and Krohn 2003; Davies and Tanner 2003; Li 1999; Nagin and Waldfogel 1995). Further, those who were incarcerated have been found to have worse physical health outcomes relative to those who were not incarcerated (Schnittker and John 2007). Additional analyses indicated that it was the presence of incarceration and not the length of incarceration that led to these adverse health outcomes.

One of the critiques of labeling theory is that those who recidivate following their initial encounters with the criminal justice system are simply at a higher risk beforehand of being a chronic offender, and it is not because they are taking on a criminal identity (Paternoster and Iovanni 1989). In other words, they have a higher criminal propensity. Further, as Hirschfield (2008) pointed out, criminal justice sanctions may not mean the same thing to everyone. In his interviews with inner-city youth, Hirschfield found that arrests did not carry a negative stigma, making it unlikely that a negative labeling effect would arise from this encounter (though it may still increase the likelihood of developing a criminal identity; see also Anderson 1999). These studies point to the need for controlling individual criminal propensity and other contextual factors when assessing the impact of criminal justice sanctions on later offending.

While traditional regression models can statistically control for these factors, analysts run the risk of over-specifying their model when they introduce too many control variables. Thus, when one wants to control for a large number of variables, propensity score matching (PSM) is a viable alternative in an attempt to control for these potential selection effects (Loughran et al. 2009; Murray et al. 2014; Nieuwbeerta et al. 2009). For instance, using data from the Pathways to Desistance study, Loughran et al. (2009) estimated the effect of institutional placement (compared to probation) on the rate of future rearrest, and whether there were any differential effects according to length of stay. Through the use of PSM, their results indicated a null placement effect on both the future rate of rearrest and the future rate of self-reported offending. As well, their findings did not uncover any marginal benefit in terms of reduced offending for an additional length of stay.

However, others have reported results that support labeling theory's arguments. Nieuwbeerta and his colleagues (2009) utilized both group-based trajectory modeling (GBTM) and PSM in an analysis of the effects of imprisonment on subsequent offending. Consistent with labeling theory, they found that being imprisoned for the first time increased the likelihood of later offending within the first three years post-release, though they were only able to match offenders with low to moderate criminal propensities. Murray and his colleagues (2014) also used PSM to assess the impact of a juvenile conviction (age 15–18), a young adult conviction (age 19–26), and an early incarceration (age 15–26) on later life success among the males in the Cambridge Study. The ages 32 and 48 life success scales comprised of several measures, including satisfactory accommodation history, satisfactory employment history, no substance use, no convictions in the prior five years, and satisfactory mental health (see also Farrington et al. 2006).[1] The results indicated that those who were convicted early in

[1] The life success scale Murray and his colleagues (2014) used included the mental well-being measure used in the current study (the General Health Questionnaire). However, the researchers did not assess the impact of criminal justice involvement on each of the items within the scale. Thus, it is unknown how criminal involvement affects the males' mental well-being specifically.

life were more likely to demonstrate poor life success later, particularly if they were convicted as a juvenile.

To the best of our knowledge, there have been three studies that have utilized a true experimental or natural experimental design to assess the impact of sanctions on later offending (Green and Winik 2010; Klein 1986; Loeffler 2013). Generally, the results do not support labeling theory's propositions. Klein (1986) reported the results of a true experiment that randomly assigned juveniles to either be released, receive community treatment, or receive a court petition following their arrest. While those who received a petition were significantly more likely to be rearrested than the other groups, the groups did not differ in their self-reported delinquent acts. Both Green and Winik (2010) and Loeffler (2013) took advantage of a natural experiment and analyzed the likelihood of recidivism among offenders who were randomly assigned to court judges. Being sentenced to prison relative to probation did not appear to have any significant effect on the likelihood of recidivism.

As a whole, this set of studies demonstrates very little clear consensus on the impact of criminal justice sanctions on later life outcomes. Further, the results from any particular study may, to some degree, rely upon the type of research design utilized. While there are exceptions, the research highlighted above tends to suggest that studies that are weaker in internal validity, such as non-experimental designs, show that criminal justice involvement has a detrimental impact on a variety of different outcomes later in life. Moreover, none of the findings emerging from true experiments offered complete support for labeling theory. While this does not necessarily imply that labeling theory's assertions are incorrect (particularly given the difficulty in conducting true experiments to test this particular theory), it does highlight the importance of considering a study's design when trying to draw causal conclusions. Next, we turn to the effects of criminal justice involvement on a specific outcome, mental health, which has been somewhat less investigated and is the focus of our current study.

Convictions and mental health

One of the major arguments of labeling theory is that criminal justice sanctions are anticipated to have adverse effects across multiple domains of a person's life. As Lemert (1967) argued, receiving the label of "offender" serves to cut a person off from several prosocial opportunities. Evidence consistent with his argument has been found in several studies, including, in particular, research on the adverse effects of incarceration. Incarceration has been shown to damage an individual's later prospects of employment in two different ways (Schnittker and John 2007). First, while incarcerated, an inmate cannot maintain the necessary credentials such as social capital, marketable skills, or a history of work. Second, following the time spent incarcerated, the mark of "felon" has been found to significantly decrease the probability of gainful employment (Pager 2003; Uggen et al. 2014). Additionally, the stigma attached to a conviction has also been found to damage social bonds and the individual's ability to form and maintain healthy relationships (Braman 2004; Hagan and Dinovitzer 1999). These consequences of being labeled, namely experiencing social isolation, unemployment, and poverty, are known to be strong predictors of poor physical and mental health outcomes (House et al. 1988; Robert and House 2000; Williams and Collins 1995).

A small body of recent research has provided support for the argument that being convicted and incarcerated may negatively affect one's mental health later in life. For instance, Turney et al. (2012) found that being incarcerated increased the risk of major depression among a sample of incarcerated fathers. Lanctot et al. (2007) compared later life outcomes among two samples: a group of institutionalized juvenile offenders and a group of juveniles who remained at home. While self-reported juvenile delinquency was found to have stronger effects on antisocial behavior when the sample was in their late twenties, being institutionalized was found to negatively impact later mental health, as indicated by scores on a measure of depression. These juveniles were also more likely to experience financial hardships, unstable employment, and lower quality relationships. These effects held after controlling for prior self-reported delinquency, suggesting that the effects were at least in part attributable to being involved in the juvenile justice system.

Piquero and his colleagues (2010) reported similar findings based on their study of later life failure among different offending trajectory groups in the Cambridge Study in Delinquent Development. These groups were identified through group-based trajectory models based on men's official conviction records from ages 10 to 40. A composite measure of life failure at ages 32 and 48 was used, combining a large variety of factors such as accommodation history, cohabitation history, employment history, substance use, self-reported offending, and mental health. The results indicated that those who followed a more serious, chronic offending trajectory exhibited significantly higher levels of life failure at both ages 32 and 48 than those who had less serious offending trajectories, even after controlling for early risk factors. However, subsequent analyses did not show significant differences in later mental health specifically, suggesting that there were no substantial differences between these trajectory groups in their emotional health.[2]

In sum, there is mixed evidence regarding the effects of justice system involvement on later mental health, with some studies suggesting that incarceration specifically increases the likelihood of depression (Lanctot et al. 2007; Turney et al. 2012), while one reported no differences among a variety of different offending trajectories (Piquero et al. 2010). Moreover, all of these studies relied upon non-experimental designs, thus limiting the extent to which causal claims can be made. However, as discussed, many potential consequences of justice system involvement such as lowered education and employment attainment have also been reported to increase the likelihood of poor mental health outcomes. As a whole, this demonstrates the importance of investigating how specific conviction and incarceration experiences impact one's later mental health using a stronger methodological approach. In the next section, we expand upon this specific issue.

[2] It is worth noting that this finding differs from that reached by Lanctot et al. (2007), which may be explained by their different analytical foci. The analyses by Lanctot et al. (2007) compared institutionalized youth to non-institutionalized youth while Piquero et al. (2010) assessed for differences between groups of offenders. As Paternoster and Iovanni (1989) pointed out, there may be a leveling-off effect of the conviction label such that, with each subsequent conviction or label applied to an individual, they have less to lose than they did upon their first conviction. This argument would suggest that we may see stronger effects when comparing individuals who were not involved in the justice system to those who experienced some involvement, as opposed to comparing differences between subgroups of offenders who were all convicted, but just at different rates.

The use of propensity score matching in analyzing effects of justice system involvement

As noted above, several prior studies have used PSM to estimate the impact of criminal sanctions on different outcomes. Recall that one of the criticisms of labeling theory is that those who recidivate following their initial justice system encounter are simply more likely in advance to be a chronic offender (Paternoster and Iovanni 1989). Thus, it is crucial to take into account an individual's criminal propensity and other potentially relevant factors when investigating the effects of sanctions on later outcomes. The PSM approach allows researchers to do so by approximating the effects of randomization to obtain less biased results. As we are usually unable to randomly assign participants to our "treatment" of criminal justice involvement, quasi-experimental methods such as PSM represent one viable option for better answering causal questions. While PSM still relies upon a knowledge-based approach (Weisburd and Hinkle 2012), it is able to take into account more potential covariates than traditional regression models without over-specifying the model.

In essence, PSM allows researchers to approximate the impact of the "treatment," or the criminal sanction in this context, on their outcome of interest while controlling, as well as possible, for potential selection effects (Apel and Sweeten 2010; Shadish et al. 2002). While randomization of the treatment condition in an experiment attempts to make it as certain as possible that the treatment is independent of the outcome, the PSM method seeks to accomplish this by matching cases that received the treatment to those who did not on the probability of receiving the treatment. After the matching procedure is completed, a researcher can conclude that the two groups are statistically equivalent on all potential confounding variables that are included in the procedure.

Current study

Labeling theory highlights the possibility that criminal justice sanctions may not deter but instead may exacerbate offending (Lemert 1967). In support of this argument, a substantial body of empirical evidence suggests that involvement in the criminal justice system has potentially serious, long-term effects across several life domains, including education, employment, and physical health (Farrington 1977; Li 1999; Murray et al. 2014; Schnittker and John 2007; Sweeten 2006). However, a smaller number of other studies including those relying upon stronger research designs have reported mixed or null effects (Green and Winik 2010; Klein 1986; Loeffler 2013). This inconsistency suggests that some of the findings in support of the theory may be distorted as they did not adequately control for the possibility of selection effects. Thus, it demonstrates the importance of relying upon as strong a research design as possible when assessing causal hypotheses.

Further, there is a small body of research that has linked criminal justice involvement to later adverse mental health outcomes, though this evidence is also mixed (Lanctot et al. 2007; Piquero et al. 2010; Turney et al. 2012). It is this gap that the current study seeks to address. Herein, we examine the impact of being convicted and incarcerated early in life on later mental health at age 32. In order to control for potential selection effects, an important issue in analyses assessing labeling theory more generally (see

Loughran et al. 2015), PSM is used to reduce the possibility of spuriousness and control for pre-existing risk factors. Based on the arguments of labeling theory and the work presented above, we test the hypothesis that those who were convicted or incarcerated early in life will be more likely to have worse mental health later in life.

Methods

The current study used data from the Cambridge Study in Delinquent Development (CSDD) to assess the impact of conviction and incarceration on later mental health. The CSDD is a prospective longitudinal study composed of a sample of 411 boys from a South London working class area (Farrington et al. 2013; Piquero et al. 2007). Almost 90% of the boys were White and British and most were born in 1953. Each male was interviewed several times between ages 8 and 18, with more recent follow-ups occurring at ages 32 and 48. The boys' parents and teachers were also interviewed early in the study. Official criminal records have been obtained for each boy starting at age 10 (the minimum age of criminal responsibility in England) to age 56. The study has a low attrition rate, as 93% of those who were still alive were interviewed in the age 48 follow-up.

Dependent variable: mental health

The key outcome variable of interest is the male's later mental health. Specifically, in the age 32 interview, the men were given the General Health Questionnaire (GHQ; Goldberg 1978). This 30-item screening questionnaire was used to identify men with potential non-psychotic psychiatric disorders, such as anxiety or depression. Each question asked the male to indicate the extent to which they suffered from a specific issue, such as an inability to concentrate or struggling with day-to-day responsibilities. Those who indicated that they were experiencing such an issue were coded as "1" and those who did not were coded as "0." The 30 items were then summed into a scale where higher scores indicated poorer mental health.

Independent variable: conviction and incarceration

In order to assess the robustness of the impact of criminal justice involvement on an individual's later mental health, three different operationalizations were utilized. All three of these measures were based on official records. The first relied upon a measure that indicated if the male received his first conviction between ages 15 and 18 ("1" indicating yes, "0" indicating no). This age range was utilized to focus on the effects of delinquency specifically on later mental health. Then, in order to assess the impact of early adulthood system involvement, the second measure indicated if the male received his first conviction between ages 19 and 26 ("1" indicating yes, "0" indicating no). The final measure indicated if, among those who were convicted, they were also incarcerated between the ages of 15 and 26 ("1" indicating yes, "0" indicating no). These two age groups were combined because few respondents within the CSDD were incarcerated. These age categories were also used in order to maintain comparability with prior

research that also assessed the impact of criminal justice involvement on later life outcomes (Murray et al. 2014).

Control variables

In order to match the respondents who were involved with the criminal justice system early in life to similar others who were not, several potential conditioning variables were used. Each of these covariates was derived from a time period preceding criminal justice system involvement. Sixteen dichotomous measures originated from teacher and parent interviews, as well as from information that psychologists received from interviews with the boys when they were between ages 8 and 10. Prior to dichotomizing, each measure's distribution was assessed and the "worst" (i.e., most criminogenic) quartiles were identified. Those who scored in this worst quartile were coded as "1" to indicate they had the risk factor while those who did not were coded as "0." While we do recognize the limitation that this coding scheme may obscure smaller differences between respondents, these same dichotomous risk factors have been used in prior research that also utilized PSM (Craig et al. 2015; Murray et al. 2014). These measures encompass several domains, such as individual-, family- and school-level factors. Specifically, individual-level factors included if the respondent had an antisocial personality, exhibited risk-taking or daring behavior, had a low verbal IQ, exhibited troublesome behavior at school, was restless or lacked concentration in class, had low grades, or was not popular with his peers. Family background factors included the presence of poor childrearing, if at least one of the parents was convicted, if the boy came from a large family, if the family lived in poor housing conditions, if they had low family income, if the boy had experienced a parent-child separation and poor parental supervision, and if the family had low socioeconomic status. A final measure indicated if the boy's school had a high delinquency rate.

These covariates were selected because prior research has identified them as important risk factors in the CSDD for future antisocial and delinquent behavior (Farrington 2003; Piquero et al. 2007). Several theories have also linked these factors to future criminal behavior as well. For example, the general theory of crime argues that inconsistent discipline and poor parental supervision lead to poor self-control (Gottfredson and Hirschi 1990). Being surrounded by delinquent peers at school may increase one's likelihood of associating with delinquent peers, a key component of social learning theory (Burgess and Akers 1966). In her developmental taxonomy, Moffitt (1993) also mentions several of these factors as being salient in the development of serious and persistent antisocial behavior, such as having a low verbal IQ, exhibiting risk-taking behavior, and having an antisocial personality.

Finally, an additional covariate representing the male's self-reported delinquency at age 14 was also used to control for the male's prior delinquent behavior. At the age 14 interview, the males were asked to indicate if they had ever committed eight specific criminal acts. These offenses included burglary, theft of motor vehicles, theft from motor vehicles, shoplifting, theft from machines, assault, drug use, and vandalism. Each criminal act was scored 0–3 to indicate its frequency (0 = never, 1 = once or twice, 2 = sometimes, 3 = frequently) and the sum from these scores was used for the age 14 delinquency score. This measure has been also used in prior research to control for prior delinquency (Murray et al. 2014).

Analytical plan

As previously mentioned, PSM was used to estimate the effects of a juvenile and young adult conviction and first-time incarceration on an individual's later mental health. This method matched each respondent who received the treatment (in this case either conviction or incarceration) to at least one other respondent who did not encounter the criminal justice system through age 26. First, three logistic regressions were used to estimate the propensity scores with the measures of conviction (ages 15–18, ages 19–26) and incarceration (ages 15–26) as the treatment variables and the control variables as the predictors (see also Craig et al. 2015; Murray et al. 2014). These propensity scores estimate the probability of each respondent receiving the "treatment" of criminal justice involvement.

The nearest-neighbor algorithm was used in matching those with criminal justice involvement to those without such involvement on the probability of involvement. A matching ratio of 2:1 was used (i.e., two untreated cases to every treated case) due to the low number of respondents who had criminal justice involvement. The standard caliper of .05 was also used. In other words, the untreated case had to have a propensity score that was within ± .05 of the calculated propensity score in order for the treated case to be matched. We then compared the balance of the covariates using a series of *t* tests between the untreated and treated groups both before and after matching in order to assess the success of the matching procedure. Additionally, we also considered the standardized differences to ensure that they were less than the absolute value of 20 (Rosenbaum and Rubin 1985). Finally, the treatment effects of each of the three forms of criminal justice involvement were analyzed using *t* tests that compared the mean of those who were involved in the criminal justice system to the mean of those who were not on their later mental health.

Results

Descriptive statistics

At the age 32 interview, the mean GHQ scale score was 4.47 (standard deviation = 4.03). Fifty-three boys were first convicted between ages 15 and 18. This subsample will be compared to those who were not convicted through their 18th year ($n = 310$). Thirty-one boys were first convicted between ages 19 and 26 and they will be compared to those who were not convicted through age 26 ($n = 279$). Finally, 34 boys were convicted and incarcerated for the first time between ages 15 and 26. These boys will be compared to those who were convicted between ages 15 and 26 but not incarcerated during this period ($n = 87$).

Propensity score matching: age 32 mental health

As discussed above, we first estimated three logistic regressions with each of the criminal justice involvement types as the outcome and the control variables as the predictors (results available upon request). Figure 1 presents the predicted probabilities based on these logistic regressions, indicating that there were cases that could be

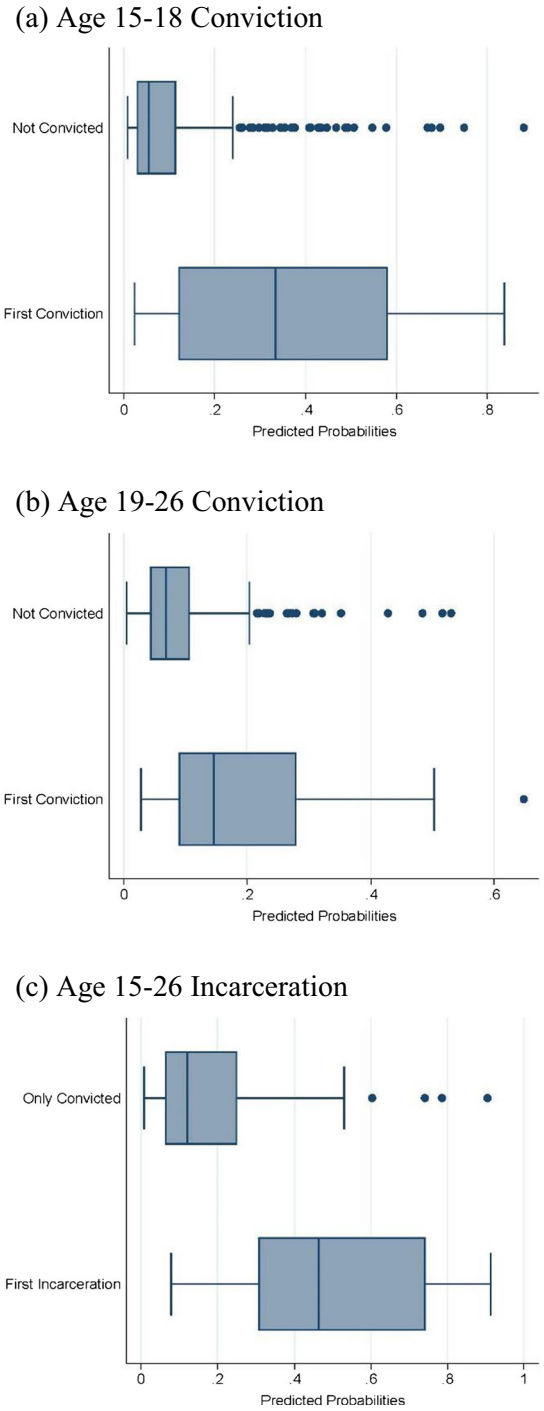

Fig. 1 Predicted probabilities for age 32 mental health. **a** Age 15–18 conviction. **b** Age 19–26 conviction. **c** Age 15–26 incarceration

matched between each of the groups. For example, as seen in Fig. 1a, while most of the males who were not convicted had a low predicted probability of receiving their first conviction between ages 15 and 18, there were several outliers who had higher probabilities of this occurring. These overlap with the majority of the convicted males who had a moderate probability of receiving their first conviction during this time period.

The data were then randomly sorted in order to avoid biasing the matching based on the original order of the data. Next, a series of *t* tests examined the differences between those who had criminal justice involvement and those who did not. The first set of columns in Tables 1, 2, and 3 presents these unmatched results. The results from the previously estimated logistic regression models were used to match those with criminal justice involvement to those without such involvement but had a similar probability of being convicted or incarcerated (caliper = .05; ratio, 2:1 with replacement).

All but one of the males who were convicted for the first time between ages 15 and 18 ($n = 52$ of 53) were able to be matched to a similarly situated non-convicted male. While all of those who did not receive their first conviction during this time could be matched ($n = 310$), due to the matching procedures that were used (2:1 matching ratio, with replacement), 60 untreated cases were successfully matched. Of the 31 males who received their first conviction between ages 19 and 26, 30 of them were able to be matched with 46 of the available 279 males who did not receive their first conviction during this time. Finally, among the 34 males who were convicted and incarcerated for the first time between ages 15 and 26, 32 of them were successfully matched with 33 of the 87 of those who had been convicted but were not incarcerated during this time.

As indicated by the second set of columns in Tables 1, 2, and 3, all post-matching covariate differences were no longer significant. Further, as the standardized differences for most covariates were below the absolute value of 20, we can assume these groups were successfully balanced in their propensity to encounter the criminal justice system. Across all three models, the standardized differences were greater than 20 for the following variables: poor childrearing, low family SES, high daring, troublesome behavior, and age 14 delinquency. Regardless, as the differences between these variables did not attain conventional statistical significance (either pre- or post-matching), this is not a large concern.

Assessment of the effect of criminal justice involvement on age 32 mental health

The prevalence and the results of the *t* tests that assessed these differences in later mental health among those with and without criminal justice involvement pre- and post-matching can be seen in Table 4.[3] The unmatched results indicate the mean differences in age 32 mental health between those who were involved with the justice system and those who were not involved in the justice system without controlling for or matching on any potential confounding variable. For instance, as can be seen in Table 4(A), while the mean mental health score among those who received their first conviction between the ages of 15 and 18 was 4.72, it was 4.41 among those who were

[3] The psmatch2 command in STATA automatically calculates the group differences as a *t* test as it is comparing the mean differences of poor mental health outcomes between the two groups.

not convicted. After the matching procedures were carried out, the mean mental health score among those who received their first conviction as a juvenile was 4.52 while it was 4.70 among those who were not convicted. The results of the *t* tests before matching can be seen in the first set of columns while the results of the *t* tests after matching are presented in the second set of columns. Neither *t* value was significant, indicating that there were no statistically significant differences in later mental health between those who were convicted for the first time between ages 15 and 18 and those who were not convicted during the same time period.

Table 4(B) presents the pre- and post-matching results among those who were convicted for the first time between ages 19 and 26. Prior to the matching procedure, the mean mental health score of those who received their first conviction as young adults was 5.16 and it was 4.34 among those who were not convicted. However, after the respondents were matched based upon their propensities for being convicted, the mean mental health score for those who were convicted was 5.23 while for their counterparts it was 4.75. None of these differences were statistically significant.

Finally, Table 4(C) indicates the results among those who were incarcerated for the first time between ages 15 and 26. Before the matching procedure, the mean mental health score among those who were incarcerated was 4.32 while the mean score for those who were not incarcerated was 5.08. Following the matching procedure, the mean mental health score was 4.03 among those who were incarcerated and 5.94 among those who were not incarcerated. Once again, these differences were not statistically significant.

In sum, these findings suggest that there were no statistically significant differences in later mental health among those who encountered the criminal justice system relative to those who did not, regardless of whether that contact was in the form of a conviction or incarceration.[4]

Supplemental analyses

We also conducted a series of supplemental PSM analyses where we assessed the impact of first criminal justice system contact in adulthood using the age 48 mental health score as the outcome. The results of these PSM analyses were similar to the prior results and showed that neither conviction measure nor incarceration measure had a significant effect on age 48 mental health outcomes. These results are available upon request from the first author.

Discussion and conclusion

The impact of criminal justice system contact, whether in the form of arrest, conviction, or incarceration, is viewed by labeling theory to be counterproductive. That is, imposition of a label on an offender is anticipated to backfire and lead to more—not less—criminal activity in the future. Aside from this key hypothesis, scholars have also examined the extent to which different types of criminal justice labels have adverse

[4] Supplementary analyses were also conducted that varied the caliper (.075, .01) and matching ratio (3:1, 1:1 with and without replacement) and the results were substantively similar to those of the current analyses.

Table 1 Balance statistics among those who were first convicted at age 15–18 on age 32 mental health, pre/post-matching *t* tests (nearest-neighbor method)

	Unmatched			Matched			Percent bias reduction	Logistic model predicting conviction propensity score	
	Convicted ($n = 53$)	Not convicted ($n = 310$)	*t* value	Convicted ($n = 52$)	Not convicted ($n = 60$)	*t* value		Coefficient	SE
Antisocial personality	.45	.16	5.26*	.44	.42	.20	93.5	.98	.57
Poor childrearing	.34	.19	2.61*	.33	.35	− .16	90.3	.48	.43
Convicted parent	.58	.18	6.99*	.58	.56	.20	95.3	1.51*	.40
High daring	.58	.22	6.03*	.58	.62	− .50	87.0	1.13*	.43
Large family size	.40	.18	3.62*	.38	.39	− .10	95.5	.40	.45
Low grades	.33	.19	2.61*	.32	.34	− .20	87.6	.43	.45
Poor housing conditions	.51	.31	2.87*	.50	.54	− .39	80.3	.53	.39
Low family income	.34	.18	2.75*	.33	.29	.42	75.9	− .49	.51
Low verbal IQ	.28	.21	1.21	.29	.23	.67	20.5	− .33	.45
Low popularity	.38	.28	1.53	.39	.36	.37	65.6	.41	.41
High delinquency rate school	.33	.15	3.46*	.32	.31	.09	95.6	.59	.50
Low family SES	.26	.16	1.78	.27	.25	.22	80.6	.27	.47
Parent-child separation	.30	.18	2.20*	.31	.26	.54	61.7	− .22	.46
Low parental supervision	.35	.15	3.62*	.36	.40	− .41	80.3	.09	.46
Troublesome behavior	.34	.14	3.64*	.33	.37	− .51	75.4	− .43	.55
Lack of concentration/restless	.32	.16	2.90*	.31	.36	− .52	70.1	− .26	.52
Age 14 delinquency	12.72	9.34	4.42*	12.73	12.11	.53	81.5	.07*	.03
Constant								− 4.31*	.51

Standardized differences are below 20 for all predictors in the matched sample

*$p < 0.05$

Table 2 Balance statistics among those who were first convicted at age 19–26 on age 32 mental health, pre/post-matching *t* tests (nearest-neighbor method)

	Unmatched			Matched			Percent bias reduction	Logistic model predicting conviction propensity score	
	Convicted ($n = 31$)	Not convicted ($n = 279$)	*t* value	Convicted ($n = 30$)	Not convicted ($n = 46$)	*t* value		Coefficient	SE
Antisocial personality	.16	.16	.06	.17	.15	.17	−333.1	−1.24	.98
Poor childrearing	.17	.19	−.35	.17	.25	−.75	−215.0	−.47	.60
Convicted parent	.19	.18	.24	.17	.15	.17	−1.1	−.99	.69
High daring	.19	.22	−.33	.20	.19	.11	52.3	−.58	.64
Large family size	.29	.17	1.67	.27	.23	.29	71.8	.75	.55
Low grades	.27	.18	1.40	.25	.26	−.10	88.2	.73	.61
Poor housing conditions	.45	.30	1.80	.43	.50	−.51	56.5	.76	.45
Low family income	.22	.17	.72	.20	.20	.00	100.0	−.24	.64
Low verbal IQ	.32	.20	1.67	.30	.28	.14	86.6	.55	.52
Low popularity	.30	.28	.23	.28	.27	.04	73.0	−.48	.54
High delinquency rate school	.18	.15	.50	.15	.17	−.15	57.2	−.78	.81
Low family SES	.19	.16	.46	.20	.12	.87	−165.3	−.05	.56
Parent-child separation	.35	.16	2.87*	.33	.37	−.27	83.2	1.25*	.49
Low parental supervision	.26	.14	1.88	.24	.18	.61	47.5	.81	.59
Troublesome behavior	.23	.13	1.43	.20	.22	−.16	81.7	1.07	.77
Lack of concentration/restless	.17	.16	.13	.17	.17	.07	25.1	−.08	.70
Age 14 delinquency	10.64	9.19	1.56	10.73	10.12	.41	57.5	.06	.04
Constant								−3.33*	.54

Standardized differences are below 20 for all predictors in the matched sample, with the exception of poor childrearing and low family SES

*$p < 0.05$

Table 3 Balance statistics among those who were first incarcerated at age 15–26 on age 32 mental health, pre/post-matching *t* tests (nearest-neighbor method)

	Unmatched			Matched			Percent bias reduction	Logistic model predicting incarceration propensity score	
	Incarcerated ($n = 34$)	Not incarcerated ($n = 87$)	*t* value	Incarcerated ($n = 32$)	Not incarcerated ($n = 33$)	*t* value		Coefficient	SE
Antisocial personality	.68	.31	3.87*	.66	.66	.00	100.0	.38	.80
Poor childrearing	.42	.29	1.39	.41	.51	−.75	28.3	−.04	.60
Convicted parent	.59	.39	1.98	.59	.53	.50	68.3	.27	.57
High daring	.68	.37	3.16*	.66	.78	−1.11	59.5	1.00	.61
Large family size	.56	.32	2.44*	.59	.62	−.25	86.8	.99	.65
Low grades	.51	.28	2.54*	.48	.44	.31	83.9	.66	.57
Poor housing conditions	.50	.54	−.40	.50	.59	−.74	−133.0	−.70	.59
Low family income	.50	.30	2.10*	.50	.56	−.49	68.9	.21	.70
Low verbal IQ	.41	.33	.80	.44	.45	−.12	80.1	−.12	.56
Low popularity	.42	.37	.54	.38	.27	.99	−110.9	−.13	.54
High delinquency rate school	.41	.28	1.43	.43	.48	−.38	63.8	.38	.63
Low family SES	.38	.22	1.85	.41	.25	1.33	4.7	.62	.59
Parent-child separation	.41	.33	.81	.41	.37	.25	60.2	−.01	.54
Low parental supervision	.34	.32	.18	.36	.38	−.13	7.8	−.52	.63
Troublesome behavior	.68	.28	4.34*	.66	.77	−.96	72.7	.93	.67
Lack of concentration/restless	.47	.23	2.64*	.50	.42	.62	67.2	.01	.65
Age 14 delinquency	16.35	11.94	3.29*	15.81	14.28	.88	65.3	.06	.04
Constant								−3.58*	.80

Standardized differences are below 20 for all predictors in the matched sample, with the exception of high daring, low popularity, low family SES, troublesome behavior, and age 14 delinquency

*$p < 0.05$

Table 4 Differences in age 32 mental health pre- and post-matching

A. First conviction at age 15–18						
	Unmatched			Matched		
	Convicted	Not convicted	*t* value	Convicted	Not convicted	*t* value
Mental health	4.72	4.41	.49	4.52	4.70	− .20
B. First conviction at age 19–26						
	Unmatched			Matched		
	Convicted	Not convicted	*t* value	Convicted	Not convicted	*t* value
Mental health	5.16	4.34	1.07	5.23	4.75	.48
C. First incarceration at age 15–26						
	Unmatched			Matched		
	Incarcerated	Not incarcerated	*t* value	Incarcerated	Not incarcerated	*t* value
Mental health	4.32	5.08	− .91	4.03	5.94	− 1.45

effects in other areas of life among labeled offenders. In this study, we concern ourselves with the unintended consequence of poor mental health among a sample of London males who were convicted and/or incarcerated in the first three decades of life and whose mental health was assessed in their 30s and 40s. An important feature of our work is that we also pay close attention to potential selection bias issues by employing propensity score methods to deal with the non-randomness of criminal justice system involvement. This type of quasi-experimental method is particularly important for research assessing the impact of criminal justice sanctions as they are not easily amenable to randomized designs in the real world.

Our analyses showed that there were no significant differences in later mental health in the 30s and 40s among those males who were convicted between ages 15 and 18 or 19 and 26, when compared to a group of propensity score-matched males who did not experience a conviction during those time periods. There was also no significant difference in later mental health between those males who were incarcerated between ages 15 and 26 and their matched non-incarcerated counterparts. In short, our analyses did not uncover any adverse effects of conviction or incarceration on later mental health.

It is important to note that these results were the same both pre- and post-matching, indicating that even before controlling for a large host of potential confounding variables, there were no significant differences in later mental health between those with earlier criminal justice involvement and those without. This is in contrast to some prior literature that has tended to support a labeling effect of criminal conviction on later outcomes (Bernburg and Krohn 2003; Davies and Tanner 2003; Farrington 1977; Farrington et al. 1978; Huizinga and Henry 2008; Li 1999; Murray et al. 2014; Nagin and Waldfogel 1995; Sweeten 2006). Taken together, this may suggest that criminal conviction and incarceration tend to increase the risk of crime/employment problems but not necessarily mental health problems. These divergent findings could be viewed from the labeling perspective. Specifically, perhaps changing one's master status to "offender" is a key mechanism for increasing later criminal behavior, but this status shift does not impact one's mental health. Further, the subsequent lost opportunities and

negative appraisals could lead to poorer employment prospects, but again, these may not affect one's mental health later in life.

These findings are also of interest as incarceration has been suggested to have a detrimental impact on prisoners' mental health (Dumont et al. 2012; Fazel et al. 2016; Kupers 1999). Specifically, Turney and her colleagues (2012) found that incarceration increased the risk of major depression among their sample of incarcerated fathers. Lanctot et al. (2007) concluded that institutionalization increased the risk of depression among juvenile delinquents when compared to juvenile offenders who remained at home. The discrepancy between these results and the present findings could imply that the high rates of mental disorders among prisoners may be due to importation or selection effects (Fazel et al. 2016; Lamb and Weinberger 1998). Another possibility is that the effects of justice system involvement on mental health may be short-lived. Though this explanation was unable to be assessed here, conviction and incarceration experiences may have had a stronger short-term effect on the CSDD men's mental health outcomes that then decreased into later adulthood.

Another possibility is that, because conviction and especially more formal criminal justice punishments (such as incarceration) was not common for these males, the pangs of punishment were not severe enough to create mental health difficulties—at least not from the perspective of the offenders and the mental health ratings used herein.[5] The particular mental health measure relied upon the General Health Questionnaire (GHQ; Goldberg 1978) which is a screening instrument to indicate if someone may be suffering from a non-psychotic minor psychiatric disorder such as anxiety or depression. Of particular relevance to the current study, it assesses the respondent's current state and whether it is similar to their "normal" state, thus potentially measuring only short-term mental health issues as opposed to longer-term conditions. This potential explanation may be useful for interpreting our findings, as previous work with the CSDD used the same GHQ measure and reported no significant differences in mental health among those with a variety of offending trajectories among the CSDD (see Piquero et al. 2010).

A final potential explanation for our findings is simply that the support for labeling theory may not be as strong as the collective body of literature suggests. As previously discussed, when one considers the internal validity of each of the studies that have assessed this theory, it is clear that studies that offer support for labeling theory *tend to be* non-experimental in nature and thus weaker in their ability to make causal claims. On the other hand, studies that utilize experimental or quasi-experimental methods offer more limited evidence for labeling theory predictions (though for exceptions, see Murray et al. 2014; Nieuwbeerta et al. 2009). In sum, the likelihood of finding detrimental effects of a criminal justice sanction on later outcomes may be weaker once researchers are better able to control for all confounding variables. While the results from this study suggest that criminal justice involvement through early adulthood does not appear to affect the likelihood of non-psychotic mental health disorders

[5] The average sentence lengths for these men were fairly short; the average time served was 1.4 years and only three offenders served longer than 3.5 years (Farrington et al. 2013).

in middle adulthood, the extent to which this is the case for other potential outcomes remains to be assessed using designs strong in internal validity.

Our study needs to be qualified on a few fronts. For example, the relatively small number of convictions—and especially incarcerations—underpowers our analysis by some extent. And while the longitudinal nature of our data and the use of propensity score matching serve as two important strengths of our work, the results need to be qualified given our small sample size.[6] On a related note, while we were interested in the impact of an individual's first criminal justice system contact on later mental health, we were unable to consider how having multiple contacts may have differential effects. Most of the crime measures in the CSDD are highly (positively) skewed as it was not very common for a male to have multiple convictions (especially early in their lives) and very few males ever experienced multiple incarcerations (Piquero et al. 2007). Third, our measure of mental health may be better suited towards assessing shorter-term mental health issues than longer-term conditions.[7] Subsequent research should consider alternative mental health measures that assess longer-term and/or more serious mental health issues. Fourth, as is the case with the majority of empirical studies on labeling theory, our criminal justice experience measures were objective indicators of whether the subject was convicted and/or incarcerated. Thus, we did not have any subjective measure regarding the extent to which the study males adopted the label and/or how they felt regarding the criminal justice involvement that they experienced.

Additionally, it would also be telling to assess if there are differences in the types and degrees of sanctions as prior research has indicated that the severity of sanctions is salient in labeling effects (Gatti et al. 2009). Fifth, the results from the propensity score models may be biased if important covariates were not included (Shadish 2013). As many of the covariates were measured several years prior to the time period covering our conviction and incarceration measures, other more proximal factors may be more salient, such as exposure to delinquent peers as an older adolescent. Attending to these and other issues will help to continue adding to the knowledge base regarding the extent to which labeling theory's hypotheses regarding unintended consequences of criminal justice system involvement hold true.

Despite these limitations, this study was the first to apply PSM to the criminal sanction–mental health relationship between early life and middle adulthood. While prior research has investigated the relationship between justice system involvement and later mental health, prior studies were unable to control for potential selection effects regarding prior differences between those who receive a criminal sanction and those who do not. Thus, the current study's findings offer evidence that early justice system involvement does not appear to have an adverse effect on later mental health in mid-adulthood. Unless randomization is possible, subsequent research should utilize PSM and other quasi-experimental approaches to approximate randomization so we can better assess causal hypotheses derived from criminological theories.

[6] The statistical power estimates for the three analyses were .06, .08, and .43 respectively.

[7] To be sure, the lack of difference in mental health outcomes prior to matching suggests that there is probably limited association between the early crime measures and later mental health outcomes. As an anonymous reviewer noted, differences may instead emerge when like groups are compared.

References

Anderson, E. (1999). *Code of the Street*. New York: Norton.

Apel, R. J., & Sweeten, G. (2010). Propensity score matching in criminology and criminal justice. In A. R. Piquero & D. Weisburd (Eds.), *Handbook of quantitative criminology* (pp. 543–562). New York: Springer.

Becker, H. S. (1963). *Outsiders*. New York: Free Press.

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

Bhati, A. S., & Piquero, A. R. (2007). Estimating the impact of incarceration on subsequent offending trajectories: deterrent, criminogenic, or null effect? *The Journal of Criminal Law and Criminology, 98*, 207–253.

Braman, D. (2004). *Doing time on the outside: incarceration and family life in urban America*. Ann Arbor: University of Michigan Press.

Burgess, R. L., & Akers, R. L. (1966). A differential association-reinforcement theory of criminal behavior. *Social Problems, 14*(2), 128–147.

Craig, J. M., Morris, R. G., Piquero, A. R., & Farrington, D. P. (2015). Heavy drinking ensnares adolescents into crime in early adulthood. *Journal of Criminal Justice, 43*(2), 142–151.

Davies, S., & Tanner, J. (2003). The long arm of the law: effects of labeling on employment. *The Sociological Quarterly, 44*, 385–404.

Dumont, D. M., Brockmann, B., Dickman, S., Alexander, N., & Rich, J. D. (2012). Public health and the epidemic of incarceration. *Annual Review of Public Health, 33*, 325–339.

Durlauf, S. N., & Nagin, D. S. (2010). The deterrent effect of imprisonment. In P. J. Cook, J. Ludwig, & J. McCrary (Eds.), *Controlling crime: strategies and tradeoffs* (pp. 43–94). Chicago: University of Chicago Press.

Farrington, D. P. (1977). The effects of public labeling. *British Journal of Criminology, 17*, 112–125.

Farrington, D. P. (2003). Key results from the first forty years of the Cambridge Study in Delinquent Development. In T. P. Thornberry & M. D. Krohn (Eds.), *Taking stock of delinquency: an overview of findings from contemporary longitudinal studies* (pp. 137–183). New York: Kluwer/Plenum.

Farrington, D. P., & Murray, J. (Eds.). (2014). *Labeling theory: empirical tests*. New Brunswick: Transaction Publishers.

Farrington, D. P., Osborn, S. G., & West, D. J. (1978). The persistence of labeling effects. *British Journal of Criminology, 18*, 277–284.

Farrington, D. P., Coid, J. W., Harnett, L., Joliffe, D., Soteriou, N., Turner, R., & West, D. J. (2006). *Criminal careers up to age 50 and life success up to age 48: new findings from The Cambridge Study in Delinquent Development (Home Office Research Study No. 299)*. London: Home Office.

Farrington, D. P., Piquero, A. R., & Jennings, W. G. (2013). *Offending from childhood to late middle age: recent results from the Cambridge Study in Delinquent Development*. New York: Springer.

Fazel, S., Hayes, A. J., Bartellas, K., Clerici, M., & Trestman, R. (2016). Mental health of prisoners: prevalence, adverse outcomes, and interventions. *The Lancet: Psychiatry, 3*, 871–881.

Garfinkel, H. (1956). Conditions of successful degradation ceremonies. *American Journal of Sociology, 61*, 420–424.

Gatti, U., Tremblay, R. E., & Vitaro, F. (2009). Iatrogenic effect of juvenile justice. *Journal of Child Psychology and Psychiatry, 50*, 991–998.

Goldberg, D. (1978). *Manual of the general health questionnaire*. Slough: NFER Nelson.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford: Stanford University Press.

Green, D. P., & Winik, D. (2010). Using random judge assignments to estimate the effects of incarceration and probation on recidivism among drug offenders. *Criminology, 48*(2), 357–387.

Hagan, J., & Dinovitzer, R. (1999). Collateral consequences of imprisonment for children, communities, and prisoners. In M. Tonry & J. Petersilia (Eds.), *Crime and justice: a review of research* (Vol. 26, pp. 121–162). Chicago: University of Chicago Press.

Hirschfield, P. J. (2008). The declining significance of delinquent labels in disadvantaged urban communities. *Sociological Forum, 23*, 575–601.

House, J. S., Landis, K. R., & Umberson, D. (1988). Social relationships and health. *Science, 241*, 540–545.

Huizinga, D., & Henry, K. L. (2008). The effect of arrest and justice system sanctions on subsequent behavior: findings from longitudinal and other studies. In A. M. Liberman (Ed.), *The long view of crime: a synthesis of longitudinal research* (pp. 220–254). New York: Springer.

Klein, M. W. (1986). Labeling theory and delinquency policy: an experimental test. *Criminal Justice and Behavior, 13*, 47–79.

Kupers, T. A. (1999). *Prison madness: the mental health crisis behind bars and what we must do about it*. San Francisco: Jossey-Bass Publishers.

Lamb, H. R., & Weinberger, L. E. (1998). Persons with severe mental illness in jails and prisons: a review. *Psychiatric Services, 49*, 483–492.

Lanctot, N., Cernkovich, S. A., & Giordano, P. G. (2007). Delinquent behavior, official delinquency, and gender: consequences for adulthood functioning and well-being. *Criminology, 45*, 131–157.

Lemert, E. M. (1967). *Human deviance, social problems, and social control*. Englewood Cliffs: Prentice-Hall.

Li, S. D. (1999). Legal sanctions and youths' status achievement: a longitudinal study. *Justice Quarterly, 16*, 377–401.

Liberman, A. M., Kirk, D. S., & Kim, K. (2014). Labeling effects of first juvenile arrests: secondary deviance and secondary sanctioning. *Criminology, 52*(3), 345–370.

Loeffler, C. E. (2013). Does imprisonment alter the life course? Evidence on crime and employment from a natural experiment. *Criminology, 51*(1), 137–166.

Loughran, T. A., Mulvey, E. P., Schubert, C. A., Fagan, J., Piquero, A. R., & Losoya, S. H. (2009). Estimating a dose-response relationship between length of stay and future recidivism in serious juvenile offenders. *Criminology, 47*(3), 699–740.

Loughran, T. A., Wilson, T., Nagin, D. S., & Piquero, A. R. (2015). Evolutionary regression? Assessing the problem of hidden biases in criminal justice applications using propensity scores. *Journal of Experimental Criminology, 11*(4), 631–652.

Matsueda, R. L. (1992). Reflected appraisals, parental labeling, and delinquency: specifying a symbolic interactionist theory. *American Journal of Sociology, 97*(6), 1577–1611.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological Review, 100*(4), 674–701.

Morris, R. G., & Piquero, A. R. (2013). For whom do sanctions deter and label?. *Justice Quarterly, 30*(5), 837–868.

Murray, J., Blokland, A., Farrington, D. P., & Theobald, D. (2014). Long-term effects of conviction and incarceration on men in the Cambridge Study in Delinquent Development. In D. P. Farrington & J. Murray (Eds.), *Labeling theory: empirical tests. Advances in criminological theory* (Vol. 18, pp. 209–235). New Brunswick: Transaction Publishers.

Nagin, D. S. (2013). Deterrence in the twenty-first century. *Crime and Justice, 42*(1), 199–263.

Nagin, D., & Waldfogel, J. (1995). The effects of criminality and conviction on the labor market status of young British offenders. *International Review of Law and Economics, 15*, 109–126.

Nieuwbeerta, P., Blokland, A., & Nagin, D. (2009). Assessing the impact of first-time imprisonment on offenders' subsequent criminal career development. *Journal of Quantitative Criminology, 25*, 227–257.

Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology, 108*, 937–975.

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

Piquero, A. R., Farrington, D. P., & Blumstein, A. (2007). *Key issues in criminal career research: new analysis of the Cambridge Study in Delinquent Development*. Cambridge: Cambridge University Press.

Piquero, A. R., Farrington, D. P., Nagin, D. A., & Moffitt, T. E. (2010). Trajectories of offending and their relation to life failure in late middle age: findings from the Cambridge Study in Delinquent Development. *Journal of Research in Crime and Delinquency, 47*, 151–173.

Robert, S. J., & House, J. S. (2000). Socioeconomic inequalities in health: an enduring sociological problem. In *Handbook of medical sociology* (5th ed., pp. 79–97). Upper Saddle River: Prentice-Hall.

Rosenbaum, P. R., & Rubin, D. B. (1985). The bias due to incomplete matching. *Biometrics, 41*, 103–116.

Schnittker, J., & John, A. (2007). Enduring stigma: the long-term effects of incarceration on health. *Journal of Health and Social Behavior, 48*, 115–130.

Shadish, W. R. (2013). Propensity score analysis: promise, reality and irrational exuberance. *Journal of Experimental Criminology, 9*(2), 129–144.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Smith, D. A., & Gartin, P. R. (1989). Specifying specific deterrence: the influence of arrest on future criminal activity. *American Sociological Review, 54*, 94–106.

Sweeten, G. (2006). Who will graduate? Disruption of high school education by arrest and court involvement. *Justice Quarterly, 23*, 462–480.

Turney, K., Wildeman, C., & Schnittker, J. (2012). As fathers and felons: explaining the effects of current and recent incarceration on major depression. *Journal of Health and Social Behavior, 53*(4), 465–481.

Uggen, C., Vuolo, M., Lageson, S., Ruhland, E., & Whitham, H. K. (2014). The edge of stigma: an experimental audit of the effects of low-level criminal records on employment. *Criminology, 52*, 627–654.

Weisburd, D., & Hinkle, J. (2012). The importance of randomized experiments in evaluating crime prevention. In B. C. Welsh & D. P. Farrington (Eds.), *The Oxford handbook of crime prevention* (pp. 446–465). New York: Oxford University Press.

Williams, D. R., & Collins, C. (1995). U.S. socioeconomic and racial differences in health: patterns and explanations. *Annual Review of Sociology, 21*, 349–386.

Jessica M. Craig is an assistant professor of criminal justice at the University of North Texas. She received both her M.S. and Ph.D. in Criminology from the University of Texas at Dallas. Her research interests include life-course/developmental criminology, delinquency, and juvenile justice. Some of her recent work has been published in the *Journal of Criminal Justice*, *Youth Violence and Juvenile Justice*, *Criminal Behaviour and Mental Health*, and *Crime & Delinquency.*

Alex R. Piquero is an Ashbel Smith Professor of Criminology and Associate Dean for Graduate Programs in the School of Economic, Political and Policy Sciences at the University of Texas at Dallas and an Adjunct Professor in the Key Centre for Ethics, Law and Governance at Griffith University. His research interests include criminal careers, criminological theory, and quantitative research methods. He currently serves as editor of *Justice Evaluation Journal*. He has received several research, teaching, and service awards and is a fellow of both the American Society of Criminology and the Academy of Criminal Justice Sciences. In 2014, he received The University of Texas System Regents' Outstanding Teaching Award and in 2018, he was named to The University of Texas System Academy of Distinguished Teachers.

Joseph Murray , PhD, is professor at the Postgraduate Programme in Epidemiology, Federal University of Pelotas, Brazil. He was previously a Senior Research Fellow at the Department of Psychiatry, Cambridge University, and at the Institute of Criminology, Cambridge University, where he obtained his M.Phil. and Ph.D., and subsequently held a British Academy postdoctoral fellowship, and a Research Fellowship at Darwin College. He was awarded the Manuel Lopez-Rey prize and the Nigel Walker prize in criminology from the University of Cambridge and the Distinguished New Scholar Award from the Division of Corrections and Sentencing, American Society of Criminology. He holds an Investigator Award from the Wellcome Trust.

David P. Farrington is an Emeritus Professor of Psychological Criminology in the Institute of Criminology, Cambridge University. In 2018, he has been chosen to receive the John Paul Scott Award of the International Society for Research on Aggression and the Herbert Bloch Award of the American Society of Criminology. He received the Stockholm Prize in Criminology in 2013. His major research is in developmental criminology, and he is the Director of the Cambridge Study in Delinquent Development, which is a prospective longitudinal survey of over 400 London males from age 8 to age 61. In addition to 756 published journal articles and book chapters on criminological and psychological topics, he has published 108 books, monographs and government publications, and 153 shorter publications (total = 1017).