This study uses propensity score matching to test the proposition that imprisonment deters future criminal activity among juvenile offenders. Using data from all court appearances of juveniles in the NSW Children's Court (Australia) between 2003 and 2004 (N = 6,196), the reoffending of a group of young offenders sentenced to control (i.e. custodial) orders (N = 376) was compared to a matched group of offenders receiving community-based sanctions. No differences were observed between the two groups. The young offenders given detention orders had a slightly lower rate of reoffending, but this difference was not significant. The results of this study indicate that, over the time period examined in this study, the imposition of a custodial sentence had no effect on the risk of reoffending.

URLs: http://researchoutput.csu.edu.au/R/-?func=dbin-jump-full&amp;object_id=33286&amp;local_base=GEN01-CSU01
http://dx.doi.org/10.1177/0004865811432585

Author Address: amcgrath@csu.edu.au

CRO Number: 33286
The effect of custodial penalties on juvenile re-offending

Andrew McGrath
School of Psychology, Charles Sturt University, NSW, Australia

Don Weatherburn
NSW Bureau of Crime Statistics and Research, NSW, Australia

**Corresponding author:** Andrew McGrath, School of Psychology, Charles Sturt University, Panorama Avenue, Bathurst NSW 2795, Australia. Email: amcgrath@csu.edu.au
Abstract

This study uses propensity score matching to test the proposition that imprisonment deters future criminal activity among juvenile offenders. Using data from all juveniles court appearances in the NSW Children’s Court (Australia) between 2003 and 2004 ($N = 6,196$), the reoffending of a group of young offenders sentenced to control (i.e. custodial) orders ($N = 376$) was compared to a matched group of offenders receiving community-based sanctions. No differences were observed between the two groups. The young offenders given detention orders had a slightly lower rate of reoffending, but this difference was not significant. The results of this study indicate that, over the time period examined in this study, the imposition of a custodial sentence had no effect on risk of reoffending.

Keywords

Juvenile offending, propensity matching, juvenile detention, recidivism, deterrence theory
Introduction

Although the extent to which punitive penal regimes enjoy public support remains unclear, imprisonment is, in most Western jurisdictions, an increasingly popular response to criminal offending (Doob and Roberts, 1988; Piquero and Steinberg, 2010). The most recent UN publication on international imprisonment records increases in prison populations between 1997 and 2007 in approximately two-thirds of countries for which data were available (Walmsley, 2010). In Australia, some 29,700 individuals were imprisoned on the national census date in 2010, representing an imprisonment rate of 172.4 individuals per 100,000 population, at an associated cost of $AU2.9 billion (Australian Bureau of Statistics, 2010; Steering Committee for the Review of Government Service Provision, 2011).

Given the costs associated with imprisonment as a sanction for criminal activity it seems reasonable to ask what we are getting in return for this investment. Three answers are commonly given to this question. The first is that prison should reduce crime by deterring the sanctioned offender. The second is that it should reduce crime by isolating, or incapacitating, the offender (Blokland and Nieuwbeerta, 2007; Piquero and Blumstein, 2007). The third is that punishment discourages crime because it acts as an expression of collective disapproval of the offender’s act. The current paper aims to investigate the first of these claims (i.e. that imprisonment acts as a specific deterrent) in a sample of juvenile offenders in NSW, Australia.

Theoretical background

Modern deterrence theory has its origins in the so-called classical criminology of Beccaria and Bentham, who argued that, among other factors, offenders would be
deterred from committing further offences by their assessment of the severity, certainty and celerity of any sanctions they might incur (Beccaria, 1764/1995; Bentham, 1823/1948). This idea was revived in Becker’s (1968) economic theory of crime. According to this theory, a person commits an offence if the expected utility of that offence exceeds the expected utility of all other legitimate alternatives (Becker, 1968: 176). The expected utility of a course of action is inversely related to the probability of apprehension and the costs associated with being caught. If Becker’s theory is true, imposing a prison sentence on persons caught engaging in criminal activity should reduce the frequency of such activity because it reduces its expected utility.

Since the revival of interest in deterrence in the 1960s, a large literature examining the potential deterrent effect of the criminal justice system has accumulated. This literature initially focused in the relationship between crime and imprisonment rates. It is now generally accepted that many of these early studies were methodologically flawed, either because they suffered from omitted variable bias or (more commonly) because they failed to take account of the reciprocal effects of crime and imprisonment on each other (Blumstein et al., 1978). Much of the research activity on deterrence in the 1980s shifted to an investigation of the deterrent effect of perceptions of sanction risk and severity (Grasmick and Bryjak, 1980; Paternoster, 1987; Piliavin et al., 1986), the role of informal sanctions (Stafford and Warr, 1993), as well as the impact of individual level factors such as low self control (Pogarsky, 2007; Tittle and Botchkovar, 2005; Wright et al., 2004). Research on sanctions risk perceptions generally provided evidence supportive of deterrence (Nagin, 1998) although the evidence appears stronger for sanction certainty than for sanction severity (Pogarsky, 2002).
Although perceptual deterrence studies provide evidence consistent with the deterrence hypothesis they do not provide direct support for the claim that offenders who are imprisoned are less likely to reoffend than offenders who receive a non-custodial sanction. There are both theoretical and empirical reasons for doubting this claim. Labelling theorists, for instance, provide compelling reasons for believing that contact with the criminal justice system increases the risk of reoffending (Becker, 1966; Braithwaite, 1989; Paternoster and Iovanni, 1989; Tannenbaum, 1957). Some maintain that the experience of imprisonment should increase subsequent offending because public identification as an offender will decrease conventional employment and educational opportunities (Bernburg and Krohn, 2003; Sampson and Laub, 1997). Others have argued that the process of labelling someone ‘criminal’ and imprisoning them is inherently criminogenic. As Becker famously put it: “The deviant is one to whom that label has successfully been applied: deviant behavior is behavior that people so label” (Becker, 1966: 9). Lemert maintained that individuals who are labelled criminal come to behave in ways consistent with that label (Lemert, 1972). Although this aspect of labelling theory has proven elusive to empirical testing, some recent studies in Australia have provided evidence that subjective perceptions of stigmatization can increase the risk of subsequent offending (McGrath, 2009; 2010).

Labelling theory is not the only theory with predictions about prison that contradict those of deterrence theory. Differential association theorists contend that prison increases the risk of reoffending because it provides an environment which reinforces deviant values and which is conducive to the acquisition of new criminal skills (Clemmer, 1940; Sykes, 1958). This theory originates in Sutherland’s proposal that crime is a result of learning that takes place during social interactions (Sutherland and Cressey, 1974). Prison, according to this theory, provides an ideal venue for such
learning. A substantial research literature exists that is generally supportive of this notion. A recent meta-analysis, for instance, found an overall effect size of .225 ($p < .001$) for differential association variables such as deviant attitudes and behavior of both peers and parents, based on a total of 385 effect sizes (Pratt et al., 2010). In summary, there appears to be substantial theoretical and empirical reasons to doubt the conclusions reached by deterrence theory in relation to the specific deterrent effect of imprisonment. We will now review the research examining this question.

**Prior research on the impact of incarceration**

Custodial penalties may exert no effect in some circumstances and a criminogenic, rehabilitative or a deterrent effect in others. The only way to determine the impact of custodial penalties is to examine their effects in a wide variety of settings. Nagin et al. (2009) recently reviewed the evidence on specific deterrence and found that the majority of studies point to a null or criminogenic effect of prison on subsequent offending. Nevertheless he and his colleagues felt unable to say with certainty what effect prison has on reoffending. Nagin et al (2009: 178) make a compelling case for the proposition that studies employing matching or regression methods should (at a minimum) control for prior record, conviction offence type, age, race and sex. Very few studies control for all of these factors, especially where juvenile reoffending is concerned.

Only two studies have used experimental methods to examine the effect of custodial penalties on juvenile reoffending. Schneider (1986) compared two groups of juvenile offenders, one of which ($n = 95$) had been randomly allocated to weekend detention, while the other of which ($n = 86$) had been allocated to a ‘restitution’ program. Rates of reoffending were found to be lower in the ‘restitution’ group, but the presiding
judge overruled the random allocation procedure in three per cent of cases allocated to custody and 11 per cent of cases allocated to restitution. This breakdown in randomization makes it hard to know what to make of Schneider’s results. In a later study, Barton and Butts (1990) found significantly higher rates of offending among juveniles randomly allocated to a home based rehabilitation program (n = 326) than among juveniles randomly allocated to a period in custody averaging about 13 months (n = 185). The results for this study were clouded by the fact that the ‘non-custody’ group actually spent an average of about six months in custody during the two year follow-up period.

Only four of the matching studies of juvenile reoffending reviewed by Nagin et al. (2009) controlled for prior record, conviction offence type, age, and sex. Three of them were conducted by Kraus (1974; 1978; 1981). Kraus (1974) compared a sample of 431 juvenile offenders given a sentence of probation with a sample of 665 juvenile offenders placed in an institution. The offenders were matched on age, age at current offence, age at time of first offence, number of previous convictions and number of previous committals to an institution. Kraus (1974) found rates of reoffending among first and recidivist offenders were generally higher for those given a detention sentence than for those given a sentence of probation. Offenders convicted of ‘take and use motor vehicle’ offences, however, committed more offences after probation than their matched counterparts did after detention. In a later, smaller study in which first-time offenders were matched on age, type and number of charges Kraus (1978) found no significant differences in re-conviction rates between custody (n = 58) and non-custody (n = 33) groups for any offence. In a third study, Kraus (1981) examined the delinquency rates among 99 juvenile offenders sent to a special school for truancy with delinquency rates among 115 juveniles placed on probation for truancy. In
addition to offence type, the two groups were matched on age, sex and prior criminal record. No differences between the two groups in rates of reoffending were observed.

The fourth matching study of custodial penalties and juvenile reoffending was conducted by Apel and Sweeten (2010a). They used data from the National Longitudinal Survey of Youth (a nationally representative sample of 8,984 youth born during the years 1980 to 1984 and living in the United States during the initial interview year in 1997) to examine the effect of incarceration on employment during the transition to adulthood. As part of this analysis they looked at the effect of incarceration on illegal income earning, a dichotomous variable coded ‘1’ if the respondent reported receiving cash for stolen items, for selling or fencing stolen property or for selling or helping to sell drugs, and ‘0’ otherwise. Two techniques were used to analyse the data; propensity score matching and fixed effects models. The propensity score matching analysis revealed no significant effect of custody on illegal income earning. The fixed effect analysis revealed a marginally significant positive effect among those placed in custody.

Only one of the regression studies of juvenile reoffending examined by Nagin et al. (2009) controlled for prior record, conviction offence type, age, race and sex. Wooldredge (1988) examined the effects of 12 different dispositions in eliminating recidivism amongst a sample of 2,038 delinquents. Five of the dispositions involved some form of detention (of varying lengths). Four of these coupled detention with some form of community supervision and treatment. The custody comparisons tested in the study were (a) case dismissed (n = 288) v one month detention (n = 31); (b) case dismissed (n = 288) v one year probation and more than one month detention (n = 18); (c) case dismissed (n = 288) v two years probation and one month detention (n =
(d) case dismissed (n = 288) vs two years probation and more than one month detention (n = 45). The follow-up periods for all offenders ranged from three to seven years. Data were analysed using both logistic regression (controlling for exposure risk) and lifetable methods but the results were the same in each case. One of the four custody comparisons was not significant at p < .05. Two were negative and significant at p < .05 (i.e. custody reduced the risk of reoffending). One was negative and borderline significant (p = .057).

The research to date on the impact of custodial penalties on juvenile reoffending suffers from a number of limitations. First, few studies have controlled for prior record, conviction offence type, age, race and sex, let alone other factors such as number of concurrent offences and age at first court appearance, that are also known to affect risk of reoffending (Chen et al., 2005; Smith and Jones, 2008). Second, as we have seen, many studies of the impact of custodial penalties on juvenile reoffending have had fairly small samples of offenders (n < 150) e.g. Kraus (1978; 1981); Schneider (1986) and Wooldredge (1988). This would have limited their power to detect an effect of custodial penalties on reoffending. Third, most studies do not appear to have taken into account time spent in custody during the follow-up period. Fourth, the results of some of the more rigorous studies are clouded by uncontrolled factors, such as the failure of randomization in Schneider’s study (1986) and the fact that custodial sanctions were combined with other penalties in three of the four custody comparisons examined by Wooldredge (1988). Finally, most of the research to date has involved American samples of juvenile offenders. The profile of those held in detention; the conditions of detention; the investment in rehabilitation and the economic and social environment into which juvenile detainees are released may vary markedly from one country to another. To the extent that they do, the impact of
custodial penalties may also vary from country to country, making it difficult to generalize results.

The purpose of this article is to report the results of a study of the effects of custodial penalties on juvenile reoffending in New South Wales (NSW), Australia. We use propensity score matching and Kaplan-Meier survival analysis (adjusted for time spent in custody) to determine whether the imposition of a custodial sentence affects the time to the next proven offence.

Propensity score matching has two main advantages over conventional regression methods when it comes to examining the effect of custodial penalties on reoffending. First, conventional regression methods do not allow the user to check whether the groups being compared are accurately matched on the covariates included in the statistical model. Second, most regression techniques assume a linear relationship between the outcome being modelled and the independent variables in the model (Apel and Sweeten, 2010b). In propensity matching, by contrast, individuals are in most cases matched non-parametrically. Propensity score matching is also superior to direct matching in that it allows researchers to match individuals on a wide range of covariates, thus avoiding the so-called ‘curse of dimensionality’ encountered by researchers using direct matching techniques (Apel and Sweeten, 2010b).

Method

Data

Data for this study comprised all finalized Children’s Court appearances in NSW in 2003 and 2004 where a juvenile was convicted of one or more charges. Where the
same individual had more than one appearance, one was chosen at random. In what follows we refer to the court appearance that resulted in entry to the study as the index court appearance. The final sample was 6,196. Information on these individuals was obtained from the reoffending database (ROD) maintained by the NSW Bureau of Crime Statistics and Research. ROD contains the records of each individual sentenced in a NSW Court since 1994. Further information regarding the matching procedures used and their accuracy can be found in Hua and Fitzgerald (2006) and Weatherburn, Lind, and Hua (2003).

**Independent variables**

The primary independent variable of interest is whether or not the juvenile offender received a custodial sentence (penalty type). In the current dataset, 376 young people received control orders. The length of sentence ranged from 2 days to 24 months ($M = 8.08$ months, $SD = 5.67$). The median sentence length was 8 months. Data on whether juveniles are remanded in custody (pre-adjudication detention) are only available at the point where the sentence is imposed. The majority of those given non-custodial sentences (85%) were bailed at the time of sentence, whereas 69.5% of those sentenced to imprisonment were in pre-detention custody. It is common for magistrates to remand juveniles in custody at their final court appearance if they are intending to impose a custodial sentence (His Honour Judge Marien, Senior Children's Court Magistrate, personal communication, 15 August, 2011).

Non-custodial sentences available to Children's Court Magistrates range from cautions to suspended control orders. In the current study, the most commonly imposed non-custodial sentence were unsupervised bonds ($N = 1,101, 17.7\%$) and fines ($N =1,030, 16.6\%)$. In addition, some 1,293 (20.8%) young people had their
cases dismissed after being cautioned. Although juvenile offenders can access drug and alcohol treatment and counselling services, few juveniles given non-custodial sentences would be placed on any formal rehabilitation program.

The additional covariates have been selected because of their established empirical and theoretical relationship with reoffending, and cover a wide range of both legal and extra-legal factors.

**Demographic**

*Gender:* Although some recent evidence suggests that the gender gap in offending is narrowing, males still make up the majority of the individuals before the justice system, and are also more likely to reoffend (Carrington, 2006; Cernkovich et al., 2008; Steffensmeier et al., 2005). Gender is thus an important control variable in any study of recidivism. In the current study, gender was coded so that 0 = girls and 1 = boys. There were 5,010 (80.9%) boys in the final sample.

*Indigenous status:* Over-representation of Indigenous Australians in the criminal justice system is a well-recognized problem (Australian Institute of Health and Welfare, 2011). Young Indigenous people are more likely to appear in court, more likely to reoffend after sentencing, and more likely to receive custodial sentences (Chen et al., 2005; Weatherburn et al., 2003). Indigenous status is thus another important control variable. This variable was coded so that 0 = non-Indigenous, and 1 = Indigenous. In the final sample, 1,319 (21.3%) individuals identified themselves as Indigenous.

*Socio-economic disadvantage:* Cross-sectional studies have shown that rates of juvenile participation in crime in NSW are much higher in areas of high economic and
social disadvantage (Weatherburn and Lind, 2001). This is to be expected because areas of high disadvantage tend to have high levels of child neglect and abuse, two factors that are known to increase the risk of juvenile involvement in crime (Weatherburn & Lind, 2001). To control for this variable we use the socio-economic indexes for areas tool (SEIFA) (Australian Bureau of Statistics, 2001), and the postcode for each offender recorded in ROD. The index score was categorized into quartiles with a higher score indicating higher levels of socio-economic disadvantage.

Remoteness: this variable is based on the Australian standard geographical classification tool (Australian Bureau of Statistics, 2008) and the postcode of residence for each offender recorded in ROD. There are five categories in this variable: 1 = living in a major city ($n = 2,860, 46.2\%$), 2 = living in inner regional areas ($n = 1,057, 17.1\%$), 3 = living in outer regional areas ($n = 1,657, 26.7\%$), 4 = living in remote areas ($n = 207, 3.3\%$), and 5 = living in very remote areas ($n = 174, 2.8\%$). Just under 4% of the sample were missing data for this variable. These individuals were given a mean value for this variable.

Legal

The positive correlation between past and future offending has been described as one of the most robust findings in criminology (Nagin and Paternoster, 2000). Theoretically, two distinct explanations for this finding have been suggested: population heterogeneity, which argues that criminal history variables act as a proxy for stable differences in individual criminal propensity, and state dependence, which asserts that contact with the criminal justice system changes the offender in such a way to increase their likelihood of recidivism (Piquero et al., 2003). It is beyond the scope of the current study to test these competing claims. Nevertheless, this literature
makes it clear that there are good theoretical and empirical reasons to include criminal history variables as covariates. In particular, measures of the following were included in the analyses:

*Age at first court appearance:* This was categorized so that 0 = first court appearance before the age of 16 \((n = 1,773, 28.6\%); 1 = first court appearance between 16 and 17 \((n = 3370, 54.4\%); and 2 = first court appearance after 17 years of age \((n = 1,053, 17\%).

*Concurrent offences:* This variable records the number of concurrent charges (both proven and unproven) at the index court appearance. This variable was categorized so that 0 = no concurrent offences \((n = 2,766, 44.6\%); 1 = 1 concurrent offence \((n = 1,553, 25.1\%); and 2 = 2 or more concurrent offences \((n = 1,877, 30.3\%).

*Counts:* This variable records the number of counts of the principle offence at the index court date. In its raw form this variable was highly skewed \((M = 1.20, SD = 1.02, range 1 – 47\). It was thus dichotomized so that 0 = one count only \((n = 5,514), and 1 = more than one count \((n = 682).

*Offence seriousness:* We considered two ways of capturing the influence of offence type. One involved creating a nominal variable that reflected the various categories of offence committed by subjects in the study. The weakness in this approach is that offenders in the sample committed a wide range of different offence types. Including a nominal variable in the model with a large number of values would have weakened the power of our analyses. The alternative was to use a measure of offence seriousness developed by MacKinnell, Poletti and Holmes (2010). This measure (known as the OSI) has two advantages. First, its values can be conveniently grouped into a small
number of categories. Second, because it is based on past sentencing practice, the OSI is a strong predictor of whether or not an offender will receive a prison sentence (MacKinnell et al., 2010). As an additional check, a logistic regression model predicting whether an individual would receive a prison sentence or not was run with a variable indicating the type of offence the person had been convicted of (instead of the OSI variable). There was no difference in model fit as judged by the AUC criterion, consequently, the offence seriousness measure was retained. Four categories of seriousness were used in the current study, where a higher score indicates a less serious offence. In category 1 (most serious) there were 3,255 individuals (52.5%); in category 2, 345 individuals (5.6%); in category 3, 1,066 individuals (17.2%); and in category 4 (least serious) there were 1,530 individuals (24.7%).

Prior convictions: This variable was also considerably skewed in its raw form, and was thus recoded so that 0 = no prior convictions \( (n = 3,461, 55.9\%) \), 1 = 1 prior conviction \( (n = 1,105, 8.4\%) \), 2 = 2 prior convictions \( (n = 520, 5.3\%) \), 3 = 3 prior convictions \( (n = 331, 3.5\%) \), 4 = 4 prior convictions \( (n = 219) \), and 5 = 5 or more previous convictions \( (n = 560, 9.0\%) \).

Prior imprisonment: This was a dichotomous variable where 0 = no prior experience of imprisonment \( (n = 5,805, 93.7\%) \), and 1 = some prior experience of imprisonment \( (n = 391, 6.3\%) \).

Prior violent offences: This was also a dichotomous variable, where 0 = no prior conviction for a violent offence \( (n = 4,689, 75.7\%) \), and 1 = a prior conviction for a violent offence \( (n = 1,507, 24.3\%) \). A violent offence was defined as aggravated assault, aggravated robbery, and aggravated sexual assault.
**Analyses**

Our object in this study is to compare the recidivism of a group of young offenders sentenced to control orders with a matched group of young offenders sentenced to community sanctions. Propensity score matching involves four discrete steps. First, the propensity score must be estimated. Second, the treatment and control groups need to be matched using this propensity score. Third, the resulting groups need to be compared to ensure that they are balanced on the covariates used (to test the conditional independence assumption). Fourth, the treatment and control groups need to be compared on the outcome variable of interest.

In the current study, the first three of these steps were conducted using psmatch2, a STATA add-in program specifically designed to perform propensity matching (Leuven and Sianesi, 2003). To perform the first step described above, the eleven covariates were regressed against a dichotomous variable indicating whether the young person had received a control order or not. Using the resulting propensity score, each member of the treated group was matched to its five nearest neighbours. The matching process was further refined by specifying a calliper weight of 0.25 and a common trim of five per cent. The first of these two constraints meant that individuals in the control group who differed from the treatment group on any covariate by more than 25 per cent of a standard deviation were omitted from the study. To test the sensitivity of this calliper size, the analyses were run again using a calliper weight of .10, however, this made little difference to the results. Specifically, the number of cases excluded did not change, and the covariate balance statistics remained essentially unchanged. The more conservative .25 calliper was therefore retained. The second constraint addresses possible violations of the common support
condition, which arise where distributions of propensity scores in the treatment and control group do not overlap (Caliendo and Kopeinig, 2008). Setting a common trim of five per cent imposes common support by dropping five per cent percent of the treatment observations at the point where the propensity score density of the control observations is the lowest. This method was originally proposed by Heckman et al (1998; 1997). Using these procedures 358 of the treatment group were matched, while a further 18 were ‘off-support’. These latter 18 individuals were omitted from all subsequent analyses.

The next step in this process is to assess the balance in baseline covariates between the control and treatment groups. There is general agreement that significance tests are an inappropriate method for assessing balance (Apel and Sweeten, 2010b; Austin, 2008). One reason for this is that typical hypothesis testing methods are affected by sample size: a change in significance level between the matched and unmatched samples might merely reflect a change in power due to the reduced sample size (Austin, 2008; Imai et al., 2008). One commonly used method to determine whether the treatment and control group are balanced is through estimation of the standardized bias (SB) (Rosenbaum and Rubin, 1985). The formula for SB is as follows:

$$SB = 100 \times \frac{\bar{x}_i - \bar{x}_{i,j}}{\sqrt{(s_i^2 + s_j^2)/2}}$$

where $s_i^2$ and $s_j^2$ are the simple standard deviations of the covariate in the treatment and control groups respectively. The standardized bias is therefore the difference in sample means divided by an estimate of the pooled standard deviation, and thus represents the difference in means between the treatment and control groups in a way
that is not dependent on the unit of measurement or the size of the sample. Although no firm criteria for determining the success of the matching procedure exists, it has been suggested that an SB below three to five per cent after matching is adequate (Caliendo and Kopeinig, 2008).

The final step in this process is to test the treatment effect. This is a less straightforward task that it may appear and has led to some controversy in the literature (Austin, 2008; Hansen, 2008; Hill, 2008; Stuart, 2008). The most common type of treatment effect is the average treatment effect (ATE), which is the average difference in outcome across the entire population. Alternatively, the average treatment on the treated (ATT) is the effect of the treatment on that part of the population who experienced the treatment. One problem with these procedures, however, is that the estimated variance of the treatment effect should also include the variance resulting from the estimation of the propensity scores as well as the imputation of common support (Caliendo and Kopeinig, 2008). Thus, relying on the resulting $t$ statistic and its associated significance level can be misleading.

An alternative procedure is survival analysis, which uses free time to reoffend as the dependent variable. This technique has a number of advantages over use of a discrete follow-up period. It is useful when follow-up periods differ, as in the current study. Time to reoffend can also be thought of as a proxy for offending frequency, which is an inherently more informative measure than a binary re-offence/no re-offence variable. The only assumption made about the underlying structure of the data is the proportional hazards assumption: that is, the model assumes the hazard rate for two individuals is parallel over time. Again, some controversy exists over whether this is an appropriate form of analyses for propensity matching studies. Austin (2008), for
instance, has argued that the log-rank test comparing the respective Kaplan-Meier survival curves of the treatment and control group is inappropriate because the matched groups are in fact not independent. In response to Austin’s comments, Hill (2008) suggested that probit regression is an appropriate alternative. Consequently, both statistics will be reported in the current study.

Results

Estimation of propensity scores

Table 1 shows the results of the logistic regression model run to estimate propensity scores. It is clear that experience of prior imprisonment is an overwhelming predictor of receiving a future detention order. Other important predictors included the number of concurrent offences, gender (with boys more likely to be given a detention order), and number of counts of the index offence. Interestingly, although Indigenous status was a strong bivariate predictor of being sentenced to detention \( OR = 2.87, p = .001 \), when included in a multivariate model with the other covariates, it became non-significant. This finding is consistent with recent research findings in Australia indicating that Indigenous offenders are sentenced using similar criteria to non-Indigenous offenders (Jeffries and Bond, 2009; Snowball and Weatherburn, 2007). The model fit statistics were robust: the Nagelkerke R-square was .345, the Hosmer and Lemeshow test \( \chi^2 = 8.86, p = .354 \), and the AUC = .902.

Table 1 about here
Matching procedure and balance tests

Matches for a total of 358 individuals in the treatment (detention) group were identified. After matching, mean sentence length was 8.03 months ($SD = 5.64$, range 2 days to 24 months), which was not significantly different from the mean sentence length of the full (i.e. $N = 376$) prison sample ($t = 0.12, p < .01$). Table 2 shows the covariate balance before and after matching. It appears that the treatment and control group were adequately balanced. SBs ranged from 0.6 (counts) to -5.70 (remoteness index). Matching subjects on their propensity scores is equivalent to matching them on the covariates used to compute those scores (Apel & Sweeten (2010b). Using the convention that an SB below three to five percent demonstrates adequate balance, we have good reason to believe that treatment and control groups in the current study were balanced on all covariates (Caliendo and Kopeinig, 2008).

Testing the effect of detention on recidivism

Having concluded that the matching procedure resulted in a treatment and control group equivalent on the observed covariates, the next step is to assess whether their likelihood of reoffending differed. Prior to matching, 34.7 per cent of the untreated group reoffended within one year of treatment, whereas 53.3 per cent of the treated group reoffended. After matching, 58.3 per cent of the control group reoffended. Figure 3 shows the survival curves for the group given detention orders and the group given community based orders in the unmatched group. It is clear that the detention group are far more likely to reoffend (Log Rank, Mantel-Cox = 92.92, $p = .001$). By contrast, Figure 4 shows the relevant survival curves for the treatment and control group in the matched sample. It is obvious here that the likelihood of reoffending is
far closer in the matched sample. In particular, the median time to reoffend for the
treatment group was 325 days, whereas for the control group it was 359 days. The
relevant log-rank test shows there is no significant difference between the groups
(Log Rank, Mantel-Cox = 1.90, \(p = .168\)). Given the criticism of this method of
analysis, an additional probit regression analysis was conducted, with recidivism
within one year of treatment as the dependent variable. As in the survival analysis
reported above, a significant difference in reoffending was observed in the unmatched
(\(SE = 0.026, z = 7.86, p < .001\)) but not in the matched (\(SE = 0.035, z = -0.85, p =
.395\)) sample. It is clear from these statistics that both groups had an equivalent
likelihood of reoffending.

Figures 1 and 2 about here

Discussion and conclusion

Theorists disagree on the effect of custodial sanctions on juvenile reoffending.
According to deterrence theory, juvenile offenders who are given a sentence of
detention should be less likely to reoffend, ceteris paribus, than juveniles given a non-
custodial sanction. Labelling theory and differential association theory predict the
opposite result. Labelling theorists would argue that, to the extent that custodial
penalties stigmatize juvenile offenders, they reinforce delinquency as a ‘master
status’. The imposition of a custodial penalty, in other words, prompts a juvenile
offender to see deviance as their defining characteristic and to act accordingly.
Differential association theory reaches the same conclusion via a different route. As
we noted in the introduction, differential association theorists maintain that prison
increases the risk of reoffending because it provides an environment which reinforces deviant values and which is conducive to the acquisition of new criminal skills.

The aim of this study was to investigate the effect of custodial penalties on juvenile offending. The results supported neither deterrence theory nor the predictions of labelling and differential association theory. The young offenders given detention sentences had a slightly lower rate of reoffending but this difference was not significant. The results of this study indicate that, over the time period examined in this study, the imposition of a custodial sentence had no effect on risk of reoffending. These results are broadly consistent with those obtained by Kraus (1974; 1978; 1981), Apel and Sweeten (2010a) and Wooldredge (1988). The present study has an advantage over these studies in that it uses a broader range of controls, a larger sample of offenders and a more sophisticated method for controlling for omitted variable bias. It is also only one of a handful of studies of the effect of custodial penalties on juvenile reoffending conducted outside North America.

This said, the present results do not by any means settle the question of whether custody has a deterrent or criminogenic effect on juvenile reoffending. Our study only followed offenders for an average of 21 months following conviction. The long-term effects of custodial penalties may be quite different to their short-term effects. Fagan and Freeman (1999) using data from a national panel study of 5,332 randomly selected youths, found that incarceration produced a significant negative effect on future employment prospects, even after adjusting for the simultaneous effects of race, human capital and intelligence. Apel and Sweeten (2010a) examined the effect of custodial penalties on employment outcomes among a sample of over a 1,000 participants in the National Longitudinal Survey of Youth. Their study is notable for
the level of control exerted over the risk of selection bias. They found that the probability of formal employment was 11 per cent lower for those given custodial penalties.

There have been no studies on the effect of juvenile detention on juvenile employment prospects in Australia, but Borland and Hunter (2000) examined the effect of an arrest record on Indigenous employment prospects using data from the 1994 National Aboriginal and Torres Strait Islander Survey. This study is significant because Indigenous juveniles make up 49.8% per cent of the NSW juvenile detention population (NSW Department of Attorney General and Justice (Juvenile Justice), 2010). Controlling for age, years completed at high school, post-school qualifications, whether the respondent had difficulty speaking English, alcohol consumption and whether the respondent was a member of the ‘stolen generation’, they found that an arrest record reduced Indigenous employment for males and females by 18.3 and 13.1 percentage points, respectively (Borland and Hunter, 2000). On this basis, Borland and Hunter (2000) estimated that differences in arrest rates for Indigenous and non-Indigenous Australians might explain about 15 percent of the difference in levels of employment between these two groups. The adverse effect of arrest and imprisonment on employment outcomes could well increase the length and depth of involvement in crime over the long term.

As with any non-experimental study of causal effects, we cannot claim to have controlled for all relevant factors. The matching procedure employed in the current study ensured that the treatment and control group were equivalent on prior record, offence seriousness, age, race and sex, race, age at first court appearance, socio-economic disadvantage, remoteness and number of counts of the principal offence.
These are all important predictors of penalty choice and recidivism. It remains possible, however, that some unmeasured difference between custody and non-custody groups biased our estimate of the effect of custodial penalties. Our only insurance against this possibility is the fact that our results are consistent with a large body of other evidence on the effect of custodial penalties on reoffending. In fact, it has been argued with increasing force in recent years that the extant research evidence suggests custodial sentences not only have no impact on reoffending, but could in fact have a criminogenic effect (Cullen et al., 2011; Nagin et al., 2009).

Some might argue that even if prison does not act as a deterrent it does help control crime through its incapacitation effect. Evidence from both international and Australian studies supports this argument so far as adults are concerned. Piquero and Blumstein, for example, have calculated that increasing the prison population by 5% will reduce crime by 1% (Piquero and Blumstein, 2007). Australian research also suggests that incapacitation may have a significant on high rate offences such as burglary (Weatherburn et al., 2006). The effectiveness of juvenile offender incapacitation in reducing crime is far less clear. Convicted juvenile offenders are far less likely to receive a custodial sentence than adult offenders. Compared with adults, the periods served in custody by juvenile offenders are also fairly short. In the current study, for instance, the mean sentence imposed was approximately 8 months. The incapacitation effect of custody on juvenile offending under current sentencing practices is therefore probably fairly small. It would be possible to increase the incapacitation effect by drastically increasing the proportion of juvenile offenders given custodial sentences and/or the length of time they spend in custody. Before we pursue this course of action, however, other options should be considered. Lipsey (1999; 2009) has identified a number of characteristics of effective interventions for
juvenile offenders. This research suggests first, that intervention programs for juvenile offenders can reduce recidivism considerably; and second, these programs are more effective when administered in the community, rather than custody. There is no doubt that incapacitation should remain part of our crime control strategy. However, it is also true that for juveniles in particular, a clear priority of juvenile justice agencies and courts should be the placement of youths into evidence-based programs.

The current results therefore strengthen the argument in favour of using custodial penalties with juvenile offenders as sparingly as possible. Aos, Miller and Drake (2006) have shown that community-based juvenile treatment/rehabilitation programs are in many cases more cost-effective than custodial penalties. Indeed, the evidence suggests that the return on investment in community-based juvenile rehabilitation programs is even greater than that for adult community-based rehabilitation/treatment programs. In the Aos et al. (2006) report just referred to, the average net present benefit (in 2006 US$) for juvenile rehabilitation programs that had a positive net present benefit was more than $22,000—more than twice the average net present benefit of adult rehabilitation programs listed in Aos et al. (2009) as having a positive net present benefit. In light of these findings it is hard to understand the continued popularity of detention as a sanction for juvenile offending. Yet in New South Wales, the number of juveniles in detention increased by 52 per cent between 2005/6 and 2009/10. It is to be hoped that the weight of evidence such as that presented here will eventually persuade policy makers to consider more cost-effective responses to the problem of juvenile crime.
Table and Figures

Table 1: Logistic regression model of the likelihood of receiving a detention order

<table>
<thead>
<tr>
<th>Covariate</th>
<th>OR</th>
<th>SE</th>
<th>z</th>
<th>P</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior imprisonment</td>
<td>6.79</td>
<td>1.15</td>
<td>11.28</td>
<td>.001</td>
<td>[4.87, 9.47]</td>
</tr>
<tr>
<td>Concurrent offences</td>
<td>2.50</td>
<td>0.21</td>
<td>10.97</td>
<td>.001</td>
<td>[2.12, 2.94]</td>
</tr>
<tr>
<td>Gender</td>
<td>2.20</td>
<td>0.49</td>
<td>3.52</td>
<td>.001</td>
<td>[1.42, 3.41]</td>
</tr>
<tr>
<td>Counts</td>
<td>1.58</td>
<td>0.26</td>
<td>2.81</td>
<td>.005</td>
<td>[1.15, 2.17]</td>
</tr>
<tr>
<td>Prior violence</td>
<td>1.53</td>
<td>0.24</td>
<td>2.68</td>
<td>.007</td>
<td>[1.12, 2.09]</td>
</tr>
<tr>
<td>Prior offences</td>
<td>1.39</td>
<td>0.06</td>
<td>7.33</td>
<td>.001</td>
<td>[1.27, 1.52]</td>
</tr>
<tr>
<td>Remoteness index</td>
<td>1.21</td>
<td>0.07</td>
<td>3.1</td>
<td>.002</td>
<td>[1.07, 1.37]</td>
</tr>
<tr>
<td>Socio-economic index</td>
<td>1.06</td>
<td>0.08</td>
<td>0.88</td>
<td>.381</td>
<td>[0.93, 1.22]</td>
</tr>
<tr>
<td>Indigenous status</td>
<td>0.96</td>
<td>0.15</td>
<td>-0.26</td>
<td>.798</td>
<td>[0.72, 1.29]</td>
</tr>
<tr>
<td>Age at first conviction</td>
<td>0.79</td>
<td>0.08</td>
<td>-2.43</td>
<td>.015</td>
<td>[0.65, 0.96]</td>
</tr>
<tr>
<td>Offence seriousness</td>
<td>0.60</td>
<td>0.04</td>
<td>-8.25</td>
<td>.001</td>
<td>[0.53, 0.68]</td>
</tr>
</tbody>
</table>

*Model $X^2 = 979$, Nagelkerke $R^2 = .345$, AUC = .902*
Table 2: Covariate balance check: standardized bias for matched and unmatched groups

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>Treated</th>
<th>Controls</th>
<th>% bias reduction in bias</th>
<th>T</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior imprisonment</td>
<td>Unmatched</td>
<td>0.457</td>
<td>0.038</td>
<td>111.2</td>
<td>35.61</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>0.430</td>
<td>0.413</td>
<td>4.6</td>
<td>95.9</td>
<td>.47</td>
</tr>
<tr>
<td>Concurrent offences</td>
<td>Unmatched</td>
<td>1.508</td>
<td>0.814</td>
<td>87.0</td>
<td>15.56</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>1.486</td>
<td>1.521</td>
<td>-4.4</td>
<td>94.9</td>
<td>.639</td>
</tr>
<tr>
<td>Gender</td>
<td>Unmatched</td>
<td>0.926</td>
<td>0.801</td>
<td>36.8</td>
<td>5.96</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>0.922</td>
<td>0.925</td>
<td>-0.8</td>
<td>97.8</td>
<td>.14</td>
</tr>
<tr>
<td>Counts</td>
<td>Unmatched</td>
<td>1.215</td>
<td>1.103</td>
<td>31.0</td>
<td>6.76</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>1.210</td>
<td>1.207</td>
<td>0.6</td>
<td>98</td>
<td>.07</td>
</tr>
<tr>
<td>Prior violence</td>
<td>Unmatched</td>
<td>0.657</td>
<td>0.216</td>
<td>99.0</td>
<td>19.9</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>0.642</td>
<td>0.628</td>
<td>3.1</td>
<td>96.8</td>
<td>.39</td>
</tr>
<tr>
<td>Prior offences</td>
<td>Unmatched</td>
<td>3.258</td>
<td>0.960</td>
<td>135.4</td>
<td>28.41</td>
<td>.698</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>3.187</td>
<td>3.154</td>
<td>1.9</td>
<td>98.6</td>
<td>.23</td>
</tr>
<tr>
<td>Remoteness index</td>
<td>Unmatched</td>
<td>2.216</td>
<td>1.938</td>
<td>26.1</td>
<td>4.96</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>2.196</td>
<td>2.257</td>
<td>-5.7</td>
<td>78.2</td>
<td>.72</td>
</tr>
<tr>
<td>Socio-economic index</td>
<td>Unmatched</td>
<td>1.954</td>
<td>2.048</td>
<td>-10.1</td>
<td>-1.84</td>
<td>.066</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>1.957</td>
<td>1.905</td>
<td>5.6</td>
<td>44.6</td>
<td>.78</td>
</tr>
<tr>
<td>Indigenous status</td>
<td>Unmatched</td>
<td>0.418</td>
<td>0.200</td>
<td>48.5</td>
<td>10.08</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>0.408</td>
<td>0.404</td>
<td>0.9</td>
<td>98.2</td>
<td>.11</td>
</tr>
<tr>
<td>Age at first conviction</td>
<td>Unmatched</td>
<td>0.923</td>
<td>0.881</td>
<td>6.4</td>
<td>1.18</td>
<td>.240</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>0.925</td>
<td>0.947</td>
<td>3.5</td>
<td>44.9</td>
<td>.46</td>
</tr>
<tr>
<td>Offence seriousness</td>
<td>Unmatched</td>
<td>1.436</td>
<td>2.186</td>
<td>-65.2</td>
<td>-11.03</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Matched</td>
<td>1.458</td>
<td>1.434</td>
<td>2.1</td>
<td>96.8</td>
<td>.33</td>
</tr>
</tbody>
</table>
Figure 1: Survival curves for the detention and community groups, unmatched sample

Figure 2: Survival curves for the detention and community groups, matched sample
References


---

1 The groups were not matched on race
2 ‘Free’ is used in this context because time in custody was subtracted from total time to reoffend