OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 0305–9049
doi: 10.1111/obes.12393

Delinquency, Arrest and Early School Leaving*

SHANNON WARD,† JENNY WILLIAMS,† and JAN C. VAN OURS‡,§,††,‡‡

†*Department of Economics, University of Melbourne, Melbourne, Victoria 3010, Australia (e-mail: shannonpmward@gmail.com), (e-mail: jenny.williams@unimelb.edu.au)*
‡*Erasmus School of Economics, Rotterdam, 3062 PA, The Netherlands*
§*Department of Economics, University of Melbourne, Melbourne, Australia*
††*Tinbergen Institute, Rotterdam, The Netherlands*
‡‡*CEPR, London, UK (e-mail: vanours@ese.eur.nl)*

Abstract

Boys typically initiate delinquent behaviour during their teenage years, and many go on to be arrested. We show that engaging in delinquency and being arrested in youth are each associated with early school leaving. The effect of delinquency on school leaving is largely driven by crimes that produce a monetary return, and the increase in school leaving is greater when onset of these types of crime, and arrest, occur at younger ages. The sizeable impact of delinquency on school leaving highlights the need for crime prevention efforts to extend beyond youth who come into contact with the justice system.

I. Introduction

Boys typically initiate delinquent behaviour during their teenage years, when they are in middle school or starting high school. Arrest first occurs at older ages and is less common. Nonetheless, youthful interactions with the justice system are far from harmless, with several recent studies demonstrating that arrest for delinquent behaviour leads to early school leaving, adversely impacting on both high school completion and college attendance (Hjalmarsson, 2008; Webbink *et al.*, 2013; Aizer and Doyle, 2015; Rud *et al.*, 2018).[1] These findings have important policy implications given the central role of education in lifetime economic well-being, as well as the increased likelihood of future arrest and imprisonment that results from early school leaving (Becker, 1962; Lochner and Moretti, 2004; Oreopoulos and Salvanes, 2011; Merlo and Wolpin, 2015; Buonanno and Leonida, 2009; Machin,

JEL Classification numbers: C4, I2, K4, D0.
*The authors are grateful to the Department of Economics, University of Melbourne for supporting this research. We thank Stephen Machin, Olivier Marie, Sarmistha Pal and other participants in the Surrey-UGPN Conference on Youth Crime and Public Policy Interventions 2017, Kevin Staub and two anonymous referees for helpful feedback on this research.

[1]The bi-directional nature of the relationship between school attendance and criminal justice interactions is evidenced by the contributions of Jacob and Lefgren (2003); Luallen (2006); Berthelon and Kruger (2011) and Anderson (2014) who find that attending school decreases contemporaneous arrests, reported incidences of crime, and prosecutions for property crime.

Marie and Vujić, 2011, 2012). However, the focus on the impact of arrest provides only a partial view of the potential consequences of delinquent behaviour for educational outcomes. Delinquents often have poor attachment to school, raising the question of whether delinquency impacts on schooling outcomes for those who are not arrested, as well as for those who are (Siennick and Staff, 2008; Hoffmann, Erickson and Spence, 2013).

From a capital accumulation perspective of crime, delinquency may lead to early school leaving even in the absence of arrest. This is because engaging in delinquency builds criminal capital, which in turn lowers the returns to education relative to crime in adulthood (Cunha and Heckman, 2007; Lochner, 2004). Lower returns make education relatively less attractive, and an early school exit more likely. Therefore, focusing only on those who are arrested fails to account for the potential for reduced schooling and attendant worsened labour market prospects that make criminal behaviour in adulthood more likely for those delinquents who avoid arrest. Ward and Williams (2015) and Webbink *et al.* (2012) take up this issue, studying the impact of delinquency and childhood conduct disorder, respectively, on education outcomes. These studies find that early onset of delinquency and conduct disorder reduces the likelihood of graduation from high school. Neither study, however, accounts for the impact of arrest on school leaving, and therefore provide a limited perspective on the nature of the relationship between delinquency, arrest and educational attainment.

We contribute to the literature on the relationship between crime and education by studying the impact of delinquent behaviour, in addition to arrest, on school leaving. Our empirical investigation uses information collected in the National Longitudinal Survey of Youth 1997. The richness of the NLSY97 data allows us to study transitions into delinquency, arrest and school leaving. Doing so has several benefits, such as assisting in disentangling potentially bi-directional relationships and distinguishing between early vs. late onset delinquency. Each are of practical importance as *ex-ante*, delinquency is potentially a cause as well as a consequence of early school leaving, and previous studies have found that early onset delinquents are more likely to persist with their offending behaviour than late onset delinquents, making early onset delinquents a particularly interesting group to study (Webbink *et al.*, 2012; Moffitt, 1993; Loeber, 1982).

We model the relationship between transitions into delinquency, arrest and school leaving using a multivariate mixed proportional hazard (MMPH) framework. The main empirical challenge in studying the relationship between delinquency, arrest and school leaving is that unobserved characteristics, such as non-cognitive skills, that lead to early school leaving may also determine delinquency and arrest (Heckman, Stixrud and Urzua, 2006). The presence of this unobserved heterogeneity is likely to render selection into delinquency and arrest endogenous to the school leaving decision. Our MMPH framework allows us to address this issue directly. Using the discrete factor approach, we model the mixing distributions capturing unobserved heterogeneity in the transition rates into delinquency, arrest and school leaving as jointly distributed. Our baseline specification finds that both arrest and delinquency increase the transition rate out of schooling, with the magnitude of the impact of arrest being around twice that of delinquency. Our extended modelling explores heterogeneities across different types of delinquent behaviours, across age at onset of these delinquent behaviours as well as across age at first arrest. It reveals that the effect of delinquency on the school leaving rate is largely driven by initiation into types of delinquency that generate a monetary return and that initiation into income generating delinquency that occurs at younger ages has a larger impact on school leaving than initiation at older ages. Similarly, being arrested has a more deleterious impact on school leaving if it first occurs at younger ages compared to older ages.

Our contributions to the literature are threefold. First, we document the timing of onset of delinquency, arrest and school leaving using longitudinal information on a large, nationally representative sample of males living in the United States. This is significant as much of the previous research using prospective longitudinal information relies on geographic specific cohort studies, with small sample sizes, and the extent to which these studies can be generalized is unclear. Second, we provide the first estimates that gauge the relative importance of delinquency and arrest on school leaving. Our consideration of both delinquency and arrest is important because many youth who engage in delinquency are not arrested in youth. Understanding whether delinquency impacts on school leaving for those who do not come into contact with the criminal justice system, as well as for those who do, is important for the targeting of policies that seek to mitigate its impacts. Our third contribution lies in providing new insights into the types of delinquent behaviours and the timing of initiation into these behaviours, and of first arrest, that lead to early school leaving. This illuminates channels through which delinquency and arrest impact on school leaving, knowledge of which is vital for designing policies to ameliorate their effects and prevent future offending.

The rest of the paper is laid out as follows. Section II offers background information motivating our investigation of the impact of delinquent activity on school leaving. This is followed by section III, which introduces the National Longitudinal Survey of Youth 1997, provides descriptive statistics and discusses the data's key features. Section IV presents the empirical model and estimation strategy used to identify the effects of delinquency and arrest on school leaving. Section V reports our baseline results along with results from specifications that allow for heterogeneous impacts of delinquency and arrest and account for suspension from school, the findings from which inform on potential mechanisms underlying the relationships we uncover. This section also reports on an investigation into the potential for reverse causality. Section VI concludes with a discussion of our findings and associated policy implications.

II. Background

The framework for studying crime proposed by Becker (1968) and extended by Ehrlich (1973), views individuals as allocating resources to crime if the expected benefit of doing so outweighs the cost. When time is the resource to be allocated and the framework is static, the cost of time spent in crime is measured by forgone labour market earnings, essentially making crime a risky occupation. Further refinements to Ehrlich's work include contributions by Grogger (1998) and Williams and Sickles (2002) who augment the model with criminal capital and social capital respectively. In these capital accumulation models, returns to legitimate and illegitimate activities depend on their respective capital stocks accumulated in previous periods. Accounting for this leads to a dynamic model of crime in which delinquency can affect school leaving because it builds criminal capital, which increases the expected monetary returns to crime relative to education (Lochner, 2004).

Early drop out is then the consequence of falling expected relative returns to education. This is particularly relevant if the wages offered to less educated young men are low, making crime a relatively more rewarding alternative (Freeman, 1994, 1991; Machin and Meghir, 2004; Lochner, 2007).

An implication of the dynamic framework that accounts for capital stock accumulation is that engaging in income generating crime should have a stronger impact on school leaving than engaging in non-income generating crime. This is because it is the expected *monetary* return to crime relative to education that determines school leaving, and experience in income generating crime is likely to have a higher future monetary return than experience in non-income generating crime.[2] A further implication of the capital accumulation model of crime is that early initiation into income generating delinquency is expected to have a larger effect on school leaving than late initiation, as those who initiate at younger ages have more periods over which to reap the rewards of their criminal capital stock.

There are other mechanisms through which delinquency may affect educational attainment. For example, if delinquent behaviour spills over into the classroom, it may result in lowered teacher expectations, suspension or expulsion from school, and increased surveillance and monitoring of delinquent students which could be perceived as harassment. Lower teacher expectations, suspension or expulsion (leading to missed classes) and an increased level of animosity due to perceived differential treatment by teachers is likely to result in an early exit from school and hence a reduced level of educational attainment (Hjalmarsson, 2008; Kupchik, 2010; Segal, 2013). However, unlike the criminal capital accumulation mechanism, these mechanisms do not imply stronger effects on school leaving for income generating delinquency relative to non-income generating delinquency. The opposite may be true if non-income generating delinquency, such as fighting at school, leads to more severe consequences (such as expulsion or suspension) than income generating acts such as theft. Furthermore, mechanisms such as suspension or expulsion imply a greater salience of delinquency for school leaving at older ages, rather than at younger ages as predicted by the capital accumulation model.

These differences in the expected impact of delinquency by type and timing of onset provide a means for discriminating between capital accumulation and alternative mechanisms in understanding the paths via which delinquency impacts on educational attainment.

III. Data and measures

The National Longitudinal Survey of Youth 1997

Our data are taken from the National Longitudinal Survey of Youth 1997 (NLSY97), a panel study of youths residing in the United States. The first round of the NLSY97 took place in 1997 and 1998 when the 8,984 respondents were between 12 and 18 years old. Since then, survey participants have been interviewed annually. We use 13 waves of data with round 13 fielded during the years 2009 and 2010 when the respondents were aged 24 to 30. The NLSY97 includes two subsamples; the first is a cross-sectional sample representative of people living in the United States (N=6,748), while the second oversamples Hispanic

[2] Lochner (2004) notes that while his model assumes that criminal returns are financial and maximizes expected income rather than utility, it can easily be generalized to include non-monetary rewards from crime.

and black people (N=2,236). We use both subsamples for our analysis. As with previous research, our analysis focuses on males as they commit the majority of delinquent acts (Moffitt, 1993; Grogger, 1998; Williams and Sickles, 2002; Hansen, 2003). Of the 4,599 males interviewed for the NLSY97, we exclude 111 who had missing information on key variables.

Outcome variables

The outcome variables of interest for this study are the age at which an individual first leaves education, the age at which an individual initiates delinquent behaviour, and the age at which an individual is first arrested. In this section, we briefly describe each of these variables. A more detailed description of their construction is provided in the Appendix.

We construct the age at which a respondent first leaves (secondary or tertiary) education using information collected on respondents' most recent grade of school completed or the number of years spent at college, and their educational institution enrollment status at the time of each interview. If the individual leaves education prior to the first wave (when individuals are aged 12–18), we use the age at which the respondent reports leaving high school in wave 1. While the focus of the paper is on investigating the impact of delinquency and arrest on educational attainment, we also consider their impacts on failing to complete high school. As individuals complete high school at age 18 or 19 in our sample, we study high school dropout using school leaving up to age 17. For this analysis, school leaving after age 17 is treated as censored.

We measure delinquency using self-reported information.[3] In wave one, respondents were asked the age at which they had first stolen items worth more than \$50, committed other property crimes, sold drugs, destroyed property or attacked someone. Subsequent waves asked whether respondents had committed each act since the last interview. We construct the age at which an individual initiates delinquent behaviour by combining the wave one information with responses from each subsequent wave. We also use this information to construct the ages at which an individual commits his first income generating delinquent act (defined as stealing items worth more than \$50, other property crimes, and selling drugs) and his first non-income generating act (defined as destroying property or attacking others). Similarly, wave one of the NLSY97 asks individuals the age at which they were first arrested, and subsequent waves inquire as to whether they have been arrested since the last interview. This information is combined to construct the age at which an individual is arrested for the first time. Note that as the survey does not ask respondents the offence for which they are arrested, we are unable to distinguish between arrest for income or non-income generating delinquent offences.

Explanatory variables

Previous research has found that the individual-level determinants of crime and educational choice include gender, age, race, cognitive ability, religiosity, local environment, family

[3] Self-reported data have an advantage over official arrest data in that arrest data may be more a reflection of police activity than individuals' offending behaviour (Hansen, 2003). There are potential issues regarding the reliability of self-reported information. These issues are discussed in the Appendix.

background and structure, and parental education (Moffitt, 1993; Levitt and Lochner, 2001; McCord, Spatz Widom and Crowell, 2001; Hansen, 2003; Mocan and Rees, 2005; Traag, Marie and Van der Velden, 2014). The NLSY97 collects rich information on the respondents' characteristics and their family backgrounds such that we are able to control for these important determinants. In addition, the longitudinal nature of the NLSY97 allows us to construct a combination of both time-varying and time-invariant control variables.

The time-invariant controls we account for in our analysis include the following individual characteristics: race (separate indicators for black and Hispanic; non-black and non-Hispanic is the omitted category); ability (standardized CAT-ASVAB score corrected for years of schooling at the test date); an indicator for attending a private or parochial school (attending a public school is the omitted category); a set of indicators for year of birth; and an indicator for experiencing puberty before the age of 12. We control for the following family characteristics: mother was a teen at the respondent's birth; the responding parent (to the parent questionnaire administered in round 1) is very religious; parents' education (of the mother and father separately); and number of siblings (younger and older separately). We also control for the following local environment variables: separate indicators for residing in the suburbs and residing in a city (residing in a rural area is the omitted category); and a set of indicators for region of residence (South; North; West; Northeast is the comparison category). In addition, we control for *time-varying* variables; specifically, parents' presence in the household using separate indicators for mother's presence and father's presence in the household at each age. Previous research has found that living in a single parent household influences delinquent behaviour by attenuating parental supervision, as well as by reducing household resources (McCord *et al.*, 2001). In an extended analysis we also control for time-varying information on suspension from school. Detailed information on the construction of the controls is provided in the Appendix.

Descriptive analysis

Figure 1 provides information on the timing of leaving education, the onset of delinquency and the age at first arrest. Specifically, the top graph shows the hazard rates for the transition out of school and the transitions into delinquency and first arrest at each age. Hazard rates for the onset of delinquency are shown for any type of delinquency as well as for income generating and non-income generating delinquency separately. The hazard rate is defined as the probability of making a transition at each age conditional on having not previously transitioned. In terms of school leaving, the hazard rate increases steeply from age 17 (shortly before most students graduate from high school at age 18 or 19), and peaks at age 19. It increases sharply again from age 21 (just before many students start graduating from college at age 22), peaking once more at age 24. In contrast, the hazard rates for the onset of (any) delinquency and non-income generating delinquency peak earlier at age 14, the hazard rate for the onset of income generating delinquency peaks at age 15, and the hazard rate for first arrest peaks at age 18.[4] The main point to be taken from the top graph of Figure 1 is that the first peak in the hazard for school leaving occurs at an older age

[4] At age 7, roughly 8% of respondents report having engaged in delinquent behaviour. This spike is partly due to our assumption that the risk of the onset of delinquency occurs from age 7. There are 249 individuals who report engaging in delinquency prior to age 7 that are recoded to initiate delinquency at age 7. This figure is 82 for income generating

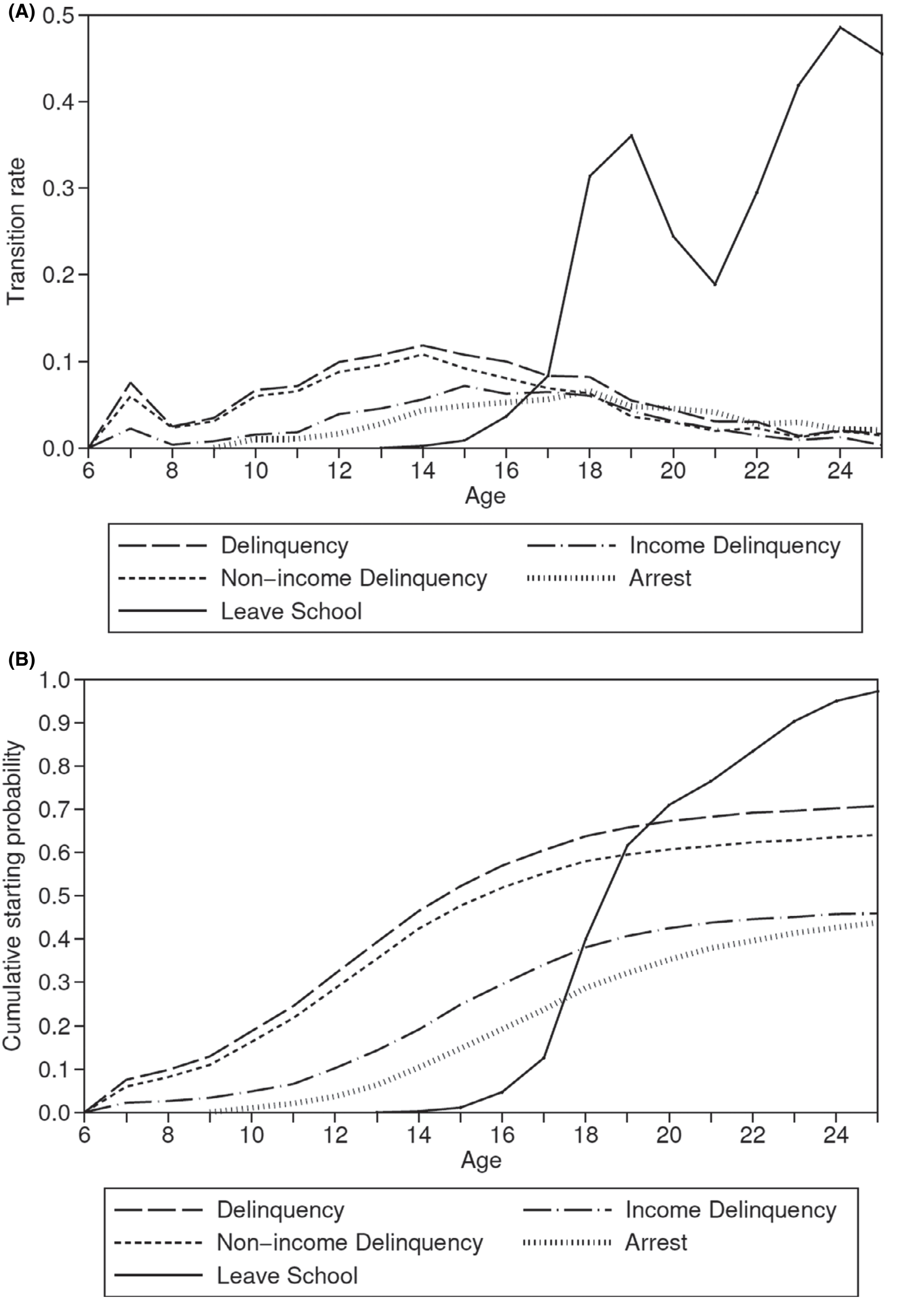

Figure 1. Transition rates and cumulative starting probabilities for first delinquency, arrest and school leaving by age

than the peak in the hazard for first arrest, which in turn occurs at an older age than the peaks for the hazard rates for the transitions into (income, non-income and any type of) delinquent behaviours.

The cumulative probability distributions for leaving education, for the onset of delinquency and for first arrest are presented in the bottom graph of Figure 1, with distributions once again presented separately for income generating, non-income generating, and any type of delinquency. The cumulative probability distributions show the proportion of the sample who have transitioned by each age. In the case of education, the distribution increases steeply over the ages of 17 through to 19, illustrating that this is when most students leave school. Specifically, the figure shows that 40% of the sample have left education by age 18. This increases to 62% by age 19, 71% by age 20, 83% by age 22 and to 97% by age 25. While the cumulative probability distribution for school leaving shows that eventually everyone will leave school, the cumulative probability distributions for first arrest and the initiation into delinquency reflect that individuals may or may not ultimately initiate delinquency or be arrested. It also shows that a higher proportion of the sample have engaged in non-income generating delinquency compared to income generating delinquency at each age, and that the probability of engaging in either increases steeply up to the age of 19. For example 25% of the sample engage in income generating delinquent activity by age 15, 38% do so by age 18, 43% by age 20, and 46% by age 25. For non-income generating delinquent activity, 48% have engaged in these acts by age 15, increasing to 58% by age 18, 61% by age 20 and 64% by age 25.[5]

Similarly, the cumulative probability distribution for first arrest shows that the probability of having been arrested by age 15 is 15%, increasing to 29% by age 18, 35% by age 20, and to 44% by age 25. From the figure, we can see that respondents are unlikely to initiate delinquent behaviour of any kind after age 19 and are unlikely to be arrested for the first time after age 25. Overall, this descriptive evidence suggests that the initiation into delinquent behaviour generally precedes school leaving, and that while it is more common for first arrest to precede school leaving, this need not be the case.[6]

Table 1 provides more detailed information on the timing of leaving education for the first time, onset of delinquency and first arrest. The first column presents the probabilities associated with the order in which the initiation of delinquency and school leaving occur. The second column does the same for the order of initiation of income generating delin-

delinquency and 188 for non-income generating delinquency (given that income and non-income generating are not mutually exclusive categories).

[5] These figures are consistent with the literature, reiterating that a substantial proportion of boys are involved in some form of delinquent behaviour during youth. Further support is provided by the Youth Risk Behavior Surveillance System (YRBSS), which shows that nationwide in the United States in 2017, 23.2% of 9th-grade males, 24.5% of 10th-grade males, 25.3% of 11th-grade males and 23.2% of 12th-grade males carried a weapon (e.g. gun, knife or club) on at least one day during the 30 days before the survey. Moreover, 33.9% of 9th-grade males, 34.7% of 10th-grade males, 25.8% of 11th-grade males and 24.1% of 12th-grade males had been in a physical fight one or more times during the 12 months before the survey. Trend analyses show these figures to have decreased during 1991-2017, suggesting that they were higher during the years the NLSY97 was fielded (Kann *et al.*, 2018).

[6] Note that the Minimum School Leaving Age (MSLA) in respondents' state of residence likely contributes to school leaving occurring after the onset of delinquency. In the year 2000, when sample members are aged 15–21, 59% of US states had a MSLA of 16, 16% had a MSLA of 17 and 24% had a MLSA of 18 according to the National Center for Education Statistics; `https://nces.ed.gov/programs/digest/d04/tables/dt04.148.asp`.

TABLE 1

First delinquency, arrest and school leaving – timing of events (percentages)

	Delinquency			
Order of events	*Any*	*Income*	*Non-income*	*Arrest*
Delinquency/arrest, school leaving	62.7	36.8	56.8	26.8
School leaving, delinquency/arrest	3.1	4.0	2.9	12.8
Delinquency/arrest = school leaving	2.1	2.5	1.8	4.1
No delinquency/arrest	32.1	56.7	38.5	56.3
Total	100.0	100.0	100.0	100.0

Note: the sample consists of 4488 observations.

quency and school leaving, the third column for the initiation of non-income generating delinquency and school leaving, and the last column for first arrest and school leaving. The first column shows that 63% of the sample initiate delinquency before leaving school, compared with only 3% who leave school prior to initiating delinquency (2% of our sample initiate delinquency and leave school at the same age and 32% are never observed to initiate delinquency).

Similarly, in the second column we see that 37% of the sample initiate income generating delinquency before leaving school, compared with only 4% who leave school before initiating income delinquency (3% do both at the same age and 57% are never observed to initiate income generating delinquency). The third column shows that 57% of the respondents initiate non-income delinquency before leaving school, as compared to 3% who leave school prior to initiating non-income delinquency (2% both leave school and initiate non-income delinquency at the same age and 39% are never observed to initiate delinquency). The last column shows that 27% of the sample are arrested before leaving school, compared with 13% who leave school prior to being arrested for the first time (4% of the sample are arrested for the first time and leave school at the same age and 56% are not observed to be arrested over the sample period).

Overall, Table 1 provides evidence that reinforces that provided by Figure 1, showing that the onset of delinquency generally precedes school leaving, and more often than not, arrest precedes school leaving for those who self-report delinquency and arrest. This provides suggestive evidence that the direction of causality linking delinquency and schooling likely runs from onset of delinquency to early school leaving, and not vice versa. With respect to arrest and schooling, there appears to be scope for both first arrest to lead to early school leaving as well as early school leaving leading to first arrest. With this in mind, our formal modelling begins by considering the impact of the onset of delinquency and first arrest on school leaving. We consider the issue of reverse causality in an extended specification.

IV. Estimation

Empirical framework

The objective of our empirical analysis is to determine: (i) the separate effects of delinquency and arrest on school leaving; and (ii) the extent to which these effects differ by the age of onset of delinquency and first arrest, and the type of offences committed, as suggested by the capital accumulation theory of crime.

Given the potential importance of early onset of delinquency and first arrest on school leaving, the empirical framework we use is the multivariate mixed proportional hazard (MMPH) model. This framework naturally permits the study of early vs. late onset in modelling the inter-relationship between transitions. In our application, the MMPH model is comprised of an equation for the transition into delinquency, an equation for the transition into arrest, and an equation for the transition out of school in which both delinquency and arrest are permitted to affect the school leaving transition rate. Transition rates are specified assuming a mixed proportional hazard function. The potential for endogenous selection into delinquency and arrest is accounted for by allowing the unobserved heterogeneity entering each of these transitions, and the unobserved heterogeneity entering into the school leaving transition rate to be correlated. Importantly, we do not impose arbitrary distributional assumptions on the joint distribution of unobserved heterogeneity (Abbring and Van den Berg, 2003; Gaure, Røed and Zhang, 2007; Heckman and Singer, 1984; Mroz, 1999). Rather we model it non-parametrically, using a discrete factor approach (Heckman and Singer, 1984; Mroz, 1999; Mroz *et al.*, 2016). This approach is particularly useful in accounting for endogeneity in the absence of instruments, and has been used effectively in many health and labour economics applications. Recent examples are Mroz *et al.* (2016); McVicar, Moschion and van Ours (2019); Moschion and van Ours (2019).[7]

Identification of the treatment effects in the MMPH model is based on the timing of events. In our application, the timing of events refers to whether, for each individual in the sample, onset of delinquency (arrest) occurs first or school leaving occurs first. Under the ‘no anticipation’ assumption, this information can be exploited via a competing risk framework to separate the causal effect of treatment from selectivity due to unobserved heterogeneity (Abbring and Van den Berg, 2003). The ‘no anticipation’ assumption ensures that cause must precede effect. While this assumption permits individuals to be forward looking in that they may have knowledge of the distribution of the timing of onset of delinquency and first arrest and act on this knowledge, it rules out information shocks about the exact future date of onset of delinquency and of first arrest that are unobserved, and hence not accounted for (Abbring and Heckman, 2007; Abbring and Van den Berg, 2003).[8] If no anticipation holds, then conditional on observed characteristics and the distribution of unobserved heterogeneity, current hazards depend only on past events and the transition processes evolve recursively. Abbring and Heckman (2007) show that under this assumption, such models are non-parametrically identified from single-spell duration data under the conditions for the identification of competing risks models based on MMPH

[7] Mroz (1999) provides Monte Carlo evaluations of this approach in simultaneous models and finds that when the true distribution of unobserved heterogeneity remains unknown and instruments are unavailable, the discrete factor approximation outperforms two-stage least squares and maximum likelihood estimators in controlling for endogeneity.

[8] More generally, the violation of the no-anticipation assumption entails that individuals: (i) have information regarding their future transition into the treatment (respectively delinquency/arrest and school leaving); (ii) know precisely when this transition will occur; (iii) alter their behaviour with respect to the other outcome of interest in anticipation (respectively school-leaving and delinquency/arrest); and (iv) that what led them to anticipate their transitions is not accounted for in the model (via observed or unobserved characteristics).

models. However, to the extent that unobserved time-varying shocks affect the onset of delinquency, first arrest and school leaving, the no anticipation assumption will not be met. For this reason, we strengthen identification through the use of exclusion restrictions on the time-varying explanatory variables. The use of exclusion restrictions is discussed in the context of the specification of the empirical model below.

The multivariate mixed proportional hazard model

Transitions out of school t_s, are modelled as a function of the onset of delinquency and first arrest. More specifically, the conditional density for leaving school at age t_s conditional on the age at which an individual first engages in delinquency t_d, the age at which he is first arrested t_a, observable characteristics x, and unobservable characteristics ε_s, is specified as

$$f_s(t_s | x, t_d, t_a, \varepsilon_s) = \theta_s(t_s | x, t_d, t_a, \varepsilon_s) \exp\left(- \int_0^{t_s} \theta_s(z_s | x, t_d, t_a, \varepsilon_s) dz_s \right) \quad (1)$$

where $\theta_s(t_s|x, t_d, t_a, \varepsilon_s)$ is the hazard rate for leaving school at age t_s and has the mixed proportional form

$$\theta_s(t_s|x, t_d, t_a, \varepsilon_s) = \lambda_s(t_s) \exp(\delta_d I(t_d < t_s) + \delta_a I(t_a < t_s) + x' \beta_s + \varepsilon_s) \quad (2)$$

Individual duration dependence is represented by $\lambda_s(t_s)$. The indicator functions, $I(t_d < t_s)$ and $I(t_a < t_s)$, equal one if the individual experiences his first delinquency before age t_s and his first arrest before age t_s, respectively, and x is a vector containing time-varying as well as time-invariant control variables, although we do not make this explicit for notational convenience.

The time-invariant controls we account for are: race (separate indicators for black and Hispanic; non-black and non-Hispanic is the omitted category); ability (standardized schooling-corrected CAT-ASVAB score); an indicator for attending a private or parochial school; a set of indicators for year of birth; an indicator for experiencing puberty before the age of 12; mother was a teen at the respondent's birth; responding parent is very religious; parents' education; number of siblings (younger and older separately); separate indicators for residing in the suburbs and residing in a city (rural residence is the omitted category); and a set of indicators for region of residence (South, North, West, Northeast is omitted). Our time-varying controls account for parents' presence in the household, with separate indicators for mother and father.

Lastly, the unobserved heterogeneity term, ε_s accounts for differences in propensities for school leaving and is assumed to be time invariant and independent of variables representing observed characteristics.[9] Those who have not left school during the observation period are assumed to have a right-censored duration until school leaving.

Note that delinquency and arrest are only permitted to impact school leaving at age t if they occur at an age prior to age t. This is consistent with delinquency and arrest capturing criminal capital stock accumulated in previous periods. It also ensures that cause precedes effect because we are unable to tell which event happened first if school leaving

[9] This is a strong assumption and, similar to the exclusion restriction in instrumental variable estimation or the time invariance of unobserved heterogeneity in fixed effects estimation, is not testable.

and delinquency, for example, occur at the same age. Similarly, the time-varying controls at age $t-1$ impact school leaving at age t. However, while the impact of delinquency and arrest are each constrained (in this specification) to have a homogeneous effect with respect to the age of first occurrence, we allow the effect of each of the time-varying covariates to be age specific. The coefficients on prior delinquency δ_d, and prior arrest δ_a, are the key parameters of interest. Positive coefficients indicate increases in the transition rate out of school, leading to lower educational attainment, whereas negative coefficients indicate decreases in school leaving, leading to higher educational attainment.

As discussed above, it is likely that correlated unobserved heterogeneity will render delinquency and arrest endogenous to transitions out of school. In order to address this endogenous selection into delinquency and arrest, we jointly model these outcomes along with school leaving. To do so, we specify the conditional density for completed durations until onset of delinquency and completed durations until first arrest, assuming mixed proportional hazards for transitions into delinquency and arrest, as follows:

$$f_j(t_j|x,\varepsilon_j)=\theta_j(t_j|x,\varepsilon_j)\exp\left(-\int_0^{t_j}\theta_j(z_j|x,\varepsilon_j)dz_j\right) \tag{3}$$

$$\theta_j(t_j|x,\varepsilon_j)=\lambda_j(t_j)\exp(x'\beta_j+\varepsilon_j) \tag{4}$$

where $j=d,a$ denotes delinquent behaviour and arrest respectively. $\lambda_j(t_j)$ accounts for duration dependence, x accounts for observed characteristics and ε_j represents unobserved characteristics. The time-invariant variables controlled for in the onset of delinquency and first arrest equations are the same as those accounted for in modelling school leaving, as are the time-varying controls, subject to the starting age of each process. Individuals who are not observed to transition into delinquency or first arrest are assumed to have right censored durations for these outcomes.

For each of the outcomes, school leaving, onset of delinquency and age at first arrest, duration dependence is modelled flexibly using a step function. For example, for school leaving, duration dependence is given by

$$\lambda_s(t_s)=\exp\left(\sum_{k=1}^{12}\lambda_{s,k}I_k(t_s)\right) \tag{5}$$

where there are 12 duration categories ($k=1,\dots,12$), 11 of which represent specific ages (14,…, 24), while the last interval is for ages greater than 24. $I_k(t_s)$ are a set of k dummy variables equal to one if the individual leaves school in duration category k, and zero otherwise. As we estimate a constant term in $x'\beta_s$, $\lambda_{s,1}$ is set to zero.

In order to account for the potential correlation in the unobserved components in the hazard rates for initiation into delinquency (ε_d) and first arrest (ε_a) and the hazard rate for school leaving (ε_s), we model them as being drawn from a discrete joint distribution, $W(\varepsilon_s,\varepsilon_d,\varepsilon_a)$. Suppose there exist two types of individual (one with high susceptibility and one with low susceptibility) for each of the outcomes, school leaving, initiation into delinquency, and first arrest.[10] This implies that the joint distribution of unobserved

[10] We experimented with alternative approaches to modelling the distribution of unobserved heterogeneity, including those described in Bonnal, Fougére and Sérandon (1997) and Mroz (1999). However, these methods imposed

heterogeneity has up to eight points of support.[11] We model these probabilities non-parametrically using a multinomial logit specification.

By integrating out the unobserved heterogeneity, we are left with the following joint density function for the duration of time until initiation into delinquency t_d, the duration of time until an individual's first arrest t_a, and the duration of time until leaving school t_s, conditional on x,

$$h_{s,a,d}(t_s,t_a,t_d|x)=\int_{\varepsilon_s}\int_{\varepsilon_a}\int_{\varepsilon_d} f_s(t_s|x,t_d,t_a,\varepsilon_s)f_a(t_a|x,\varepsilon_a)f_d(t_d|x,\varepsilon_d)dW(\varepsilon_s,\varepsilon_a,\varepsilon_d) \qquad (6)$$

As discussed above, identification of causal effects is based on the timing of events. In addition, we make use of exclusion restrictions on the time-varying variables to strengthen identification. These exclusion restrictions aid in identifying treatment effects because their *past* values affect past transition rates into delinquency and first arrest and thereby assist in revealing the distribution of unobserved heterogeneity, while only affecting the *current* transition probabilities out of education through the selection process (Gaure *et al.*, 2007).[12] Essentially, those with values of (time-varying observed) covariates indicating a high probability of an *early* transition into delinquency, for example, but who do not transition until *later*, reveal a low expected value of their (unobserved) propensity towards transitioning (Nordberg and Røed, 2009). In fact, Gaure *et al.* (2007) show that some exogenous variation in transition rates over time means no time-invariant covariates are required to achieve identification, while Brinch (2007) explains that the required variation in covariates over time is minimal.

It is important to account for the fact that our data are collected in yearly interviews, as this means we do not measure time continuously, but rather, on an annual basis. We therefore adjust our approach for time-grouped data, formulating the likelihood function such that we integrate the hazards over discrete time intervals.

V. Empirical results

Delinquency, arrest and school leaving

The key results from estimating the multivariate system of equations for initiation into delinquency, first arrest and school leaving are presented in panel *a* of Table 2. The first column reports the parameter estimates assuming that the unobserved heterogeneity terms in the delinquency, arrest and school leaving processes are independent (i.e. assuming no selection into delinquency and arrest), the second column reports key estimates assuming the unobserved heterogeneities are correlated (which accounts for endogenous selection

a lot of structure on the distribution of unobserved heterogeneity and were not flexible enough for our application, which includes the possibility of a zero transition rate (for the onset of delinquency and first arrest).

[11] We model the second mass point $(\varepsilon_{d,2},\varepsilon_{a,2},\varepsilon_{s,2})$ for each behaviour differentially, such that

$$\varepsilon_{d,1}+\varepsilon_{d,22}=\varepsilon_{d,2},\quad \varepsilon_{s,1}+\varepsilon_{s,22}=\varepsilon_{s,2},\quad \varepsilon_{a,1}+\varepsilon_{a,22}=\varepsilon_{a,2}$$

where $\varepsilon_{d,1},\varepsilon_{a,1},\varepsilon_{s,1}$ and $\varepsilon_{d,22},\varepsilon_{a,22},\varepsilon_{s,22}$ are estimated.

[12] To make matters more concrete, while mother present at age 13 may impact on the onset of delinquency at age 14, and onset of delinquency at 14 may impact school leaving at age 15, mother present at age 13 is excluded from the hazard rate for school leaving at age 15, which depends on mother present at age 14.

TABLE 2

Parameter estimates – multivariate hazard model: effects on school leaving

	Independent UH		Correlated UH	
a. Three processes				
Delinquency	0.27	(6.7)***	0.27	(5.4)***
Arrest	0.52	(11.8)***	0.46	(9.0)***
−Loglikelihood	27,292.4		27,038.8	
LR test statistic	507.2**			
b. Four processes				
Income delinquency	0.33	(7.6)***	0.34	(6.5)***
Non-income delinquency	0.08	(2.0)**	0.09	(1.7)*
Arrest	0.47	(10.5)***	0.41	(7.7)***
−Loglikelihood	34,657.0		33,850.5	
LR test statistic	1613.0**			

Notes: 4488 observations. Absolute t-statistics in parentheses; ***$P<0.01$, **$P<0.05$, *$P<0.1$ See for other Correlated UH parameter estimates Table A.2.

into delinquency and arrest). The LR test comparing both specifications confirms that the unobserved heterogeneities determining school leaving, initiation into delinquency and first arrest are not independent.[13] Nevertheless, the parameter estimates representing the effects of delinquency and arrest on school leaving are hardly affected by accounting for endogenous selection into delinquency and arrest through the introduction of correlated unobserved heterogeneity.

We find that arrest increases the school leaving rate by 58% (100(exp(0.46)-1)) and delinquency increases it by 31% (100(exp(0.27)-1)). These effects are statistically significant at the 1% level. Three significant findings arise from our baseline results. First, we corroborate previous studies finding that arrest reduces schooling. Second, we reveal that after accounting for the effect of interaction with the justice system (as reflected by arrest) on school leaving, delinquency reduces schooling. The third significant finding is that the magnitude of the impact of arrest on school leaving is around twice that of delinquency.

The full set of results, including the coefficient estimates for the explanatory variables can be found in Appendix Table A.2. This table shows that the effect of the individual-level determinants of crime and educational choice that previous research has found to be important, such as race, cognitive ability, religiosity, local environment, family background and structure, and parental education are all generally as expected (Moffitt, 1993; Levitt and Lochner, 2001; McCord *et al.*, 2001; Hansen, 2003; Mocan and Rees, 2005; Traag *et al.*, 2014).

For example, a lower, and therefore later, transition out of school is associated with being black[14], having a higher level of ability, attending private or parochial schools,

[13] The LR test statistic (507.2) exceeds the critical value of 11.1 for a χ^2_5 distribution at the 5% level of significance. The distribution of unobserved heterogeneity is presented and discussed in the appendix (see Table A.2).

[14] This result is consistent with previous studies that control for an ability (CAT-ASVAB or AFQT) score in schooling and wage regressions (Hjalmarsson, 2008; Neal and Johnson, 1996; Murnane, Willett and Levy, 1995).

having a parent who is very religious, having a mother or father with at least a high school education, living in the suburbs, having a mother present in the household at age 16 and at age 19 and older, and a father present in the household at ages 14 and older. The school leaving rate is higher for those whose mothers were teens at the respondents' births and for those with a greater number of siblings.

In terms of the transition into delinquency, the likelihood is lower for those who are Hispanic, have higher ability, attend a private or parochial school, have a father who has graduated college, a mother present in the household at ages 6, 9, 11, 13 and 16, and a father present in the household at ages 8, 17, and at age 19 and older. The likelihood of delinquency is higher for those who enter puberty early, live in the suburbs or city, and for those who live in the South. The likelihood of transitioning into first arrest is lower for those who have higher ability, attend a private or parochial school, have a very religious parent, a father who graduated college, a mother who graduated high school but not college, a mother present in the household at ages 9, 10, 13 and 14, and a father present in the household at ages 13, 15, and at age 17 and older. The likelihood of arrest is higher for those who enter puberty early, who live in the suburbs or city, and those who live in the North or South.

Income vs. non-income generating delinquency

In a model of crime that accounts for human and criminal capital accumulation, experience in delinquency builds criminal capital, which in turn lowers the expected monetary return to education relative to crime (Lochner, 2004).[15] Given that experience in income generating delinquency is likely to have a higher future monetary return than experience in non-income generating delinquency, engaging in income generating delinquency should have a stronger impact on school leaving. To investigate whether the data support this prediction, we replace the equation for the transition into delinquency with two separate equations for the transitions into income and non-income generating delinquency, and allow each to separately affect the school leaving rate. Income generating delinquent acts include stealing more than \$50 worth of goods, other property crimes, and selling drugs. Non-income generating delinquent acts include destroying property and attacking others. Combined with the hazards for first arrest and school leaving, this results in a four equation system of hazards in which the four processes have a correlated error structure, allowing for endogenous selection into arrest and each type of delinquency.

The key results from maximum likelihood estimation of the four equation system of hazards are presented in panel *b* of Table 2. The first column reports the parameter estimates assuming that the unobserved heterogeneity terms in the four equations are independent, while the second column reports key estimates assuming the unobserved heterogeneities are correlated. The parameters of interest are the coefficients on arrest, income and non-income generating delinquency in the equation for school leaving. Again, the LR test shows that the unobserved heterogeneity terms in the four processes are not independent.[16] And

[15] A lower expected return to education may lead to early school leaving among non-delinquent as well as delinquent youth. For example, Apel *et al.* (2008) find that uptake of formal employment at age 16 among NLSY97 sample members who are not working at age 15 increases the likelihood of dropping out of school at age 16.

[16] The test statistic of 1613.0 exceeds the critical value of 16.9 for a χ^2_9 distribution at the 5% level of significance. The distribution of unobserved heterogeneity is presented and discussed in the appendix (see Table A.3).

again, the main parameter estimates are not much affected by the inclusion of correlated unobserved heterogeneity that accounts for endogenous selection into delinquency and arrest.

Income generating delinquency, non-income generating delinquency and arrest all increase the rate at which individuals leave school. The coefficients are significant at the 1% level in the case of income generating delinquency and arrest. The estimates imply that the school leaving rate is increased by 51% (100(exp(0.41)-1)) following first arrest, and by 41% (100(exp(0.34)-1)) following the onset of income generating delinquency. Onset of non-income generating delinquency increases the school leaving rate by just 9% (100(exp(0.09)-1)).

In summary, we find that the effect of initiation into delinquency on school leaving is primarily driven by initiation into income generating delinquency. This is not predicted by mechanisms such as delinquency spilling over to school and impacting on teachers' expectations, or leading to suspension or expulsion. However, these results are consistent with predictions from a capital accumulation model of crime in which income generating delinquency builds criminal capital stock. The final point to take from these findings is that the magnitude of the effect of income generating delinquency on school leaving is similar to that of arrest.

Age at initiating delinquency and first arrest

The four equation system of hazards presented in Table 2 constrains the effects of initiating income and non-income generating delinquency and being arrested for the first time to have the same impact on school leaving regardless of the age at which they occur. However, the accumulation of criminal capital or salience of critical ages may lead to differential effects by age. For example, a capital accumulation model of crime predicts that earlier initiation into income generating delinquency will have a larger effect on school leaving than later initiation. Similarly, being arrested when younger is likely to have a more detrimental effect on schooling than being arrested when older. In contrast, if the mechanism via which delinquency impacts on school leaving is suspension or expulsion, then delinquency at older ages may have greater salience for school leaving compared to delinquency at younger ages.

To investigate these hypotheses, we expand the model with correlated unobserved heterogeneity in panel *b* of Table 2 by allowing the effects of (income and non-income generating) delinquency and arrest to vary according to whether they first occur at age 15 or younger (given that the hazard rates for non-income and income delinquency peak at ages 14 and 15, respectively), at age 16 or 17, or when aged 18 or older (by which time individuals have reached the age of majority). Panel *a* of Table 3 repeats the baseline results reported in panel *b* of Table 2 for ease of reference. The key results are presented in the first column of panel *b* of Table 3. An LR test comparing the specifications in panels *a* and *b* of Table 3 confirms the existence of differential age of onset effects for delinquency and arrest on school leaving, with a test statistic of 17.6 exceeding the critical value of 12.6 for a χ^2_6 distribution at the 5% level of significance.

Beginning with the parameter estimates on the arrest variables in the equation for school leaving, we find that the indicator for first arrest occurring by age 15 and the indicator for first arrest occurring at ages 16 or 17 are each statistically significant at conventional levels,

TABLE 3

Parameter estimates – multivariate hazard model: effects on school leaving – sensitivity analysis

	Baseline set-up		*Censoring*	
a. Baseline set-up				
Income delinquency	0.34	(6.5)***	0.58	(3.1)***
Non-income delinquency	0.09	(1.7)*	-0.19	(0.9)
Arrest	0.41	(7.7)***	0.53	(3.3)***
−Loglikelihood	33,850.5		27,165.9	
b. Age-specific initiation effects				
Income delinquency				
Initiate ⩽ age 15	0.35	(6.1)***	0.65	(3.3)***
Initiate at age 16 or 17 (16)	0.29	(3.6)***	0.26	(0.8)
Initiate ⩾ 18	0.30	(2.8)**		
Non-income delinquency				
Initiate ⩽ age 15	0.07	(1.3)	−0.22	(1.1)
Initiate at age 16 or 17 (16)	−0.001	(0.0)	−0.06	(0.2)
Initiate ⩾ 18	0.25	(1.9)*		
Arrest				
Initiate ⩽ age 15	0.52	(8.3)***	0.56	(3.3)***
Initiate at age 16 or 17 (16)	0.34	(4.4)***	0.44	(1.6)
Initiate ⩾ 18	0.15	(1.5)		
−Loglikelihood	33,841.7		27,164.8	
LR-test statistic	17.6		2.2	

Notes: 4488 observations. Absolute t-statistics in parentheses; ***$P < 0.01$, **$P < 0.05$, *$P < 0.1$

while the indicator for first arrest occurring at 18 or older is not. In terms of the magnitude of effects, the school leaving rate is higher for those who are first arrested earlier. Similarly, for income generating delinquency, we find that while initiation at all ages has a statistically significant effect on school leaving, the effect is largest for initiation that occurs early (by age 15).[17] Overall, the findings for both income generating delinquency and arrest are consistent with predictions from a capital accumulation model of crime, suggesting that the accumulation of criminal capital is a potential mechanism via which delinquency and arrest impacts school leaving.

In contrast, initiation into non-income generating delinquency has a statistically significant effect on school leaving only when initiation occurs when aged 18 or older. Initiation into non-income generating delinquency that occurs close to the end of high school, at age 18 or older, therefore appears to be particularly disruptive, suggesting salient or vulnerable ages in terms of the school leaving decision. This may occur if violent behaviour at school leads to suspension in the final year of high school. Youth may see little value in returning to school when their suspension ends, especially if they are unable to graduate with their class.

[17] An LR test fails to reject the null hypothesis of equal coefficients on initiating income delinquency at ages 16 or 17 and initiating income delinquency at age 18 or older (at the 5% level of significance).

TABLE 4

Parameter estimates – multivariate hazard model: effects of suspensions on delinquency, arrest and school leaving

Effects of:	*Delinquency*		*Arrest*		*School Leaving*	
Delinquency	–		–		0.23	(4.7)***
Arrest	–		–		0.41	(8.2)***
Suspended age 6–12	0.47	(4.9)***	0.85	(5.6)***	–	
Suspended age 13–15	0.45	(4.4)***	0.81	(8.7)***	0.76	(5.2)***
Suspended age 16–17	0.85	(5.0)***	0.71	(5.9)***	0.31	(4.9)***
–Loglikelihood			26934.4			

Notes: The effects of a suspension at age t on delinquency, arrest and school leaving are estimated at age $t+1$ with the following exceptions: effects on arrest are from age 10 onwards; effects on school leaving are from age 14 onward. 4488 observations. Absolute t-statistics in parentheses; $^{***}P < 0.01$

School suspensions

In order to better gauge the extent to which alternative channels play a role in school leaving, we augment our baseline model by including time-varying information on suspension from school.[18] Similar to time-varying controls for presence of parents, we include an indicator for suspended at age t in the hazards for delinquency, arrest and school leaving at age $t+1$.[19]

The key coefficient estimates from this extended specification are reported in Table 4. In terms of school leaving, the results indicate that suspension from school significantly increases school leaving at subsequent ages. However, suspension at older ages does not appear to have a larger effect on school leaving compared to suspension at earlier ages, as salience would predict. Importantly, accounting for school suspensions has little effect on the magnitude or significance of the impact of delinquency or arrest on school leaving, suggesting that these behaviours impact school leaving through channels other than suspension.

In addition to its effect on school leaving, the estimates in Table 4 show that suspension from school is a significant determinant of the onset of delinquency and first arrest (at the subsequent ages). Overall these estimates show that suspension from school is associated with worse outcomes for young males in terms of the onset of delinquency, first arrest and school leaving. Nonetheless, they do not provide support for the hypothesis that suspension from school at older ages is more strongly associated with school leaving, nor do they account for the effect of delinquency and arrest on school leaving.

The impact of initiating delinquency and first arrest on dropping out of high school

It is failing to complete high school that the literature has identified as causally related to adult crime and incarceration and, for this reason, it is of particular interest to determine

[18] In this analysis, suspension from school is treated as exogenous. While it is likely that selection into suspension from school is endogenous to school leaving, addressing this issue is beyond the scope of this study.

[19] Note that transitions into delinquency, arrest and school leaving begin at ages 7, 10 and 14 respectively. As a consequence, suspensions from ages 6, 9 and 13 enter these respective hazards.

the impact of delinquency and arrest on high school dropout specifically, rather than school leaving in general. To investigate this, and noting that most in our sample are 18 or 19 years old when they complete high school, we redefine our school leaving variable so that it is censored at age 17 for those who leave school at ages greater than 17. This allows us to focus on leaving school before completing high school.

The second column of Table 3 reports the relevant parameter estimates with the redefined (censored) school leaving variable. A comparison of the parameter estimates in panel *a* shows that when school leaving is censored at age 17, the qualitative findings remain unchanged, with income producing delinquency and arrest leading to a higher school leaving rate. However, the magnitude of the estimated effects is larger for the outcome variable measuring the school leaving rate up to age 17 (prior to high school completion). Similarly, the coefficient estimates in panel *b* show that early income generating delinquency and arrest lead to higher rates of school leaving compared to later income generating delinquency and arrest. However, the coefficient estimates in the second column are larger than those in the first column, indicating larger effects on leaving school before high school completion. All in all, this suggests that the onset of income generating crime and first arrest have large effects on dropping out of high school and these effects are greater the earlier that the onset of these behaviours occur.

Reverse causality

We now address the potential for school leaving to affect the onset of delinquency and being arrested for the first time. In doing so in a hazard setting, we note that while everyone eventually leaves school, it is failing to complete high school that has been shown to increase the risk of offending, arrest and incarceration in adulthood. In our data, individuals usually complete high school at age 18 or 19. For this reason, we distinguish between school leaving that is early (up to age 17), typical (at age 18 or 19) and late (age 20 or older). We therefore augment our baseline specification reported in panel *a* of Table 2 by including in the starting rates for delinquency and arrest an indicator function, $I(t_s \leqslant 17 < t_j)$, that is equal to one if the individual leaves school at or before age 17 and this occurs before the current period, along with similar indicators for those who leave school at ages 18 or 19, $I(18 \leqslant t_s \leqslant 19 < t_j)$, and those who leave school at age 20 or older, $I(20 \leqslant t_s < t_j)$. The hazard rates for the onset of delinquency and first arrest, therefore, become

$$\begin{aligned}\theta_j(t_j|x, t_s, \varepsilon_j) = \lambda_j(t_j)\exp(x'\beta_j + \delta_j^e I(t_s \leqslant 17 < t_j) + \delta_j^m I(18 \leqslant t_s \leqslant 19 < t_j) \\ + \delta_j^l I(20 \leqslant t_s < t_j) + \varepsilon_j)\end{aligned} \quad (7)$$

where $j = d, a$ for delinquency and arrest, respectively, and t_s is the age at which an individual leaves school.[20] The parameters δ_j^e, δ_j^m and δ_j^l measure the effects of early, typical (average) and late school leaving on initiation into delinquent behaviour and first arrest, with the comparison category being those that do not leave school prior to initiating delinquency and first arrest respectively.

[20] We only allow school leaving to affect delinquency or arrest if it happens in a previous period, as we are unable to tell which occurred first if they happen at the same age.

TABLE 5

Reverse causality: effect of leaving school at different ages on delinquency and arrest

Effects of:	*Delinquency*	*Arrest*		*School leaving*	
Delinquency	–	–		0.26	(5.4)***
Arrest	–	–		0.50	(9.7)***
School leaving $\leqslant$ age 17	0.06 (0.3)	0.98	(7.6)***	–	
School leaving at age 18 or 19	0.15 (0.8)	0.37	(3.2)***	–	
School leaving $\geqslant$ age 20	−0.22 (0.6)	0.13	(0.8)	–	
−Loglikelihood		27010.5			

Notes: 4488 observations. Absolute t-statistics in parentheses; $^{***}P < 0.01$, $^{**}P < 0.05$, $^{*}P < 0.1$.

In Table 5, we present key estimation results for the multivariate hazard model in which we allow for reverse causality in the relationships between delinquency and school leaving, and arrest and school leaving. The results reveal no evidence that early, typical or late school leaving affects the onset of delinquency, with each of the coefficients on the school leaving variables in the delinquency equation being statistically insignificant at conventional levels of significance. This is not surprising given that only 3% of sample members leave school before initiating into delinquency. And while we find that the coefficient on late school leaving (age 20 or older) is insignificant in the arrest equation, the coefficient on typical school leaving (at ages 18 or 19), and the coefficient on early school leaving (at or before age 17) are both positive and significant at the 1% per cent level. Nevertheless, our baseline results are not very sensitive to accounting for reverse causality.

Our finding that early (and to a lesser extent, typical) school leaving increases the likelihood of first arrest is consistent with previous studies reporting that higher levels of education and being enrolled in school reduces the likelihood of arrest (Lochner and Moretti, 2004; Merlo and Wolpin, 2015; Buonanno and Leonida, 2009; Anderson, 2014). We also find that early school leaving has no effect on initiation into delinquency, a result consistent with the findings of Sweeten, Bushway and Paternoster (2009).

Heterogeneity by race and socio-economic status

We investigate the potential for heterogeneous effects of delinquency and arrest across race and socio-economic status. The potential for differences across race arises if there are differential impacts of criminal justice sanctions, or differences in reporting delinquent behaviour by race. Differences across socio-economic status may arise if, for example, families with more resources are able to ameliorate the impact of criminal justice interactions on school leaving.[21]

In order to investigate whether there are differential effects of arrest and delinquency on school leaving we expand our baseline model to include, respectively, interaction terms with race (black, Hispanic, with non-Hispanic and non-black as the comparison category)

[21] It would also be interesting to investigate heterogeneity in the effects of delinquency and arrest by school performance and reason for school leaving. However, there is substantial non-response on both variables. For about one in three individuals in our sample, information about school performance and the reason for school leaving is missing.

TABLE 6

Sensitivity analysis: heterogeneity of the effects of delinquency and arrest on school leaving

	Effect Delinquency		*Effect Arrest*		*-Loglikelihood*
a. Baseline	0.27	(5.4)***	0.46	(9.0)***	27038.8
b. Effect	0.40	(6.3)***	0.44	(6.5)***	
Effect × black	−0.26	(2.6)***	0.08	(0.8)	
Effect × Hispanic	−0.34	(3.3)***	−0.01	(0.0)	27031.5
c. Effect	0.24	(4.2)***	0.53	(9.1)***	
Effect × high SES	0.06	(0.8)	−0.30	(3.3)***	27033.7

Notes: Parameter estimates multivariate hazard model. See for the baseline estimates Table 2 second column, panel *a*. Observations: 4488. Absolute t-statistics in parentheses. ***$P < 0.01$, **$P < 0.05$, *$P < 0.1$.

and with high socio-economic status, defined as having a father with more than a high school education (37% of our sample).[22] Table 6 reports the results from investigating heterogeneity by race in panel *b* and by socio-economic status in panel *c*. Our baseline results are reproduced in panel *a*.

In terms of differences across race, we note that the interaction terms between the race indicators and arrest reported in panel *b* are not individually statistically significant. This suggests a common effect of arrest across race. In the case of delinquency, however, we do find negative and statistically significant interaction terms with the indicators for race, with delinquency having the largest effect on school leaving for those who are non-black/non-Hispanic. An LR test comparing the specifications in panels *a* and *b* confirms that the race interaction terms are jointly significant.[23] As we find no evidence of differential effects of arrest by race, it seems likely that the significant differences of delinquency on school leaving by race reflects differential reporting. Given that previous research suggests that non-black/non-Hispanics are least likely to underreport delinquent activity, our baseline coefficient estimate of the effect of delinquency on school leaving may be considered a lower bound of the true effect (Hindelang, Hirschi and Weis, 1981; Lochner and Moretti, 2004).

The results of investigating differences across socio-economic status are presented in panel *c* of Table 6. An LR test comparing the specifications in panels *a* and *c* confirms that the interaction terms are jointly significant.[24] The coefficient estimates in panel *c* show that the interaction term between the high socio-economic status indicator and delinquency is not individually statistically significant. This suggests a common effect of delinquency on the school leaving rate across socio-economic status, with the school leaving rate increased by the same percentage for a member of any socio-economic status.[25] In the case of arrest, however, the interaction term with the indicator for high socio-economic status is negative

[22] If no information is available on the respondent's father's education, we substitute the respondent's mother's education (the case for 250 respondents).

[23] The test statistic of 14.6 exceeds the critical value of 9.5 for a χ^2_4 distribution at the 5% level of significance.

[24] The test statistic of 10.2 exceeds the critical value of 6 for a χ^2_2 distribution at the 5% level of significance.

[25] These results are consistent with those of Webbink *et al.* (2012) who investigate the effect of childhood conduct disorder on high school graduation and find little effect of socio-economic background.

and statistically significant, indicating that arrest has a larger effect on school leaving for those of low socio-economic status. This suggests that boys from families with more resources are better able to overcome the averse consequences on school leaving of arrest.

VI. Discussion

Our paper investigates the relationship between delinquency, arrest and early school leaving. We do so using a multivariate mixed proportional hazard framework to model the transitions into delinquency, arrest and school leaving as a system. Using nationally representative data on males from the United States, and accounting for correlated unobserved heterogeneity using a discrete joint mixing distribution, we provide evidence that initiation into delinquency as well as first arrest leads to early school leaving. When homogeneous treatment effects are imposed across types of delinquency, and on the age of initiation into delinquency and the age of first arrest, we find that being arrested for the first time has roughly twice the effect of initiation into delinquency on the likelihood that an individual leaves school. Given that twice as many males who are enrolled in school have been delinquent and not arrested compared to the number arrested, our estimates imply that the overall impact of delinquency on school leaving is at least as large as that of arrest. We also examine the extent to which reverse causality is an issue. We find that leaving school early does increase the likelihood of first arrest, but only if it occurs by age 19, with larger effects when school leaving occurs earlier. We find no evidence that early school leaving leads to initiation into delinquency.

Motivated by a capital accumulation framework, we investigate whether (i) the onset of income generating delinquency has a larger effect on school leaving than the onset of non-income generating delinquency; and (ii) whether the effect of income and non-income generating delinquency and arrest is greater for onset at younger ages than for onset at older ages. These more detailed analyses reveal that the effect of delinquency on school leaving is largely driven by initiation into delinquency that is income generating, and that early initiation into income generating delinquency has a larger effect on school leaving than later initiation. Similarly, first arrest leads to school leaving only if it occurs prior to age 18, and its effect is larger for younger ages at first arrest. These findings are not consistent with mechanisms whereby delinquency impacts on school leaving through reduced teacher expectations, or through suspension or expulsion from school, which would imply larger effects at older ages, and for non-income generating, violent offences. These findings are consistent, however, with predictions from a capital accumulation model of crime, suggesting that in addition to human capital, criminal capital accumulation may be an important mechanism through which delinquency impacts the school leaving decision. In contrast, we find that initiation into non-income generating delinquency only has an effect on school leaving if it occurs close to the end of high school, at age 18 or older. This may be evidence of mechanisms such as violent delinquency spilling over to school, leading to expulsion or suspension, as these mechanisms are particularly salient to the school leaving decision at older ages. However, our investigation of the role of school suspension on school leaving does not support this conjecture, as suspension at ages 14 and 15 are found to have a larger impact on school leaving than suspensions that occur at ages 16 or 17.

A significant policy relevant finding of this research is that focusing interventions solely on those who come to the attention of the criminal justice system would miss a large part of the vulnerable population of boys whose delinquency puts them at risk of early school leaving. And while delinquents who avoid arrest may remain undetected by law enforcement, they are likely to have come to the attention of their school teachers and principals. This suggests that school-based interventions are likely to be an effective means of reaching this group. For example, 'Becoming a Man' is a school-based prevention program aimed at improving the social-cognitive skills of disadvantaged male youths in grades 7–10 from high crime Chicago neighbourhoods. Recent research by Heller *et al.* (2013) uses a randomized control trial to evaluate the effects of this program, finding that it induces significant reductions in offending.

The second policy relevant finding from our research is that the results are consistent with predictions from a model of criminal and human capital accumulation, suggesting that criminal capital in addition to human capital accumulation may be a mechanism via which delinquency leads to early school leaving. Underlying decision making in this model are beliefs about potential returns to work and crime, conditional on accumulated capital stocks. This suggests that policies that aim to improve knowledge about the returns to education, targeted at disadvantaged individuals who may not otherwise have access to reliable information, may reduce early school leaving (Jensen, 2010; Hansen, 2003). For example, Neumark and Joyce (2001) show that students' participation in school-to-work programs (such as job shadowing, mentoring, cooperative education, and internships or apprenticeships) improves individuals' subjective probabilities of obtaining a high school diploma, and increases their perceived likelihood of future labour market activity. Similarly, career academies (which are school-based programs in which students are supported in gaining work experience, and are taught using career-related materials) have been shown to reduce dropout rates of at-risk students (Belfield and Levin, 2007; Kemple and Snipes, 2000), and programs such as the Philadelphia Futures' Sponsor-A-Scholar (SAS) program that reduce the costs of education in terms of both tuition fees and attitudes towards schooling increase individuals' expected returns to education (Heckman, 2000; Johnson, 1997).

Finally, our research highlights the complex nature of the relationship between delinquency, arrest and education. Our findings provide useful insights for policy development. Specifically, we show that there are a large group of delinquents who avoid arrest before leaving school, but whose reduced level of educational attainment is as important as that of the group that have been arrested. While we narrow our focus to consider the impacts of initiation into delinquency and first arrest on school leaving, persistence of offending is also an important dimension to consider in understanding the relationship between delinquency, arrest and schooling. Pursuing this issue is another important area for future research.

Final Manuscript Received: May 2020

References

Abbring, J. H. and Heckman, J. J. (2007). Econometric evaluation of social programs. In Heckman, J. J. and Leamer, E. E. (eds), *Handbook of Econometrics*, Vol. 6, chapter 3, Amsterdam: Elsevier, pp. 4777–5752.

Abbring, J. H. and Van den Berg, G. J. (2003). 'The nonparametric identification of treatment effects in duration models', *Econometrica*, Vol. 71, pp. 1491–1517.

Aizer, A. and Doyle Jr, J. J. (2015). 'Juvenile incarceration, human capital, and future crime: evidence from randomly assigned judges', *Quarterly Journal of Economics*, Vol. 130, pp. 759–803.

Anderson, D. M. (2014). 'In school and out of trouble? The minimum dropout age and juvenile crime', *Review of Economics and Statistics*, Vol. 96, pp. 318–331.

Apel, R., Bushway, S. D., Paternoster, R., Brame, R., and Sweeten, G. (2008). 'Using state child labor laws to identify the causal effect of youth employment on deviant behavior and academic achievement', *Journal of Quantitative Criminology*, Vol. 24, pp. 337–362.

Becker, G. S. (1962). 'Investment in human capital: a theoretical analysis', *Journal of Political Economy*, Vol. 70, pp. 9–49.

Becker, G. S. (1968). 'Crime and punishment: an economic approach', *Journal of Political Economy*, Vol. 76, pp. 169–217.

Belfield, C. R. and Levin, H. M. (2007). *The Return on Investment for Improving California's High School Graduation Rate*. California Dropout Research Project Report 2.

Berthelon, M. E. and Kruger, D. I. (2011). 'Risky behavior among youth: incapacitation effects of school on adolescent motherhood and crime in Chile', *Journal of Public Economics*, Vol. 95, pp. 41–53.

Bonnal, L., Fougére, D., and Sérandon, A. (1997). 'Evaluating the impact of French employment policies on individual labour market histories', *Review of Economic Studies*, Vol. 64, pp. 683–713.

Brinch, C. N. (2007). 'Nonparametric identification of the mixed hazards model with time-varying covariates', *Econometric Theory*, Vol. 23, pp. 349–354.

Buonanno, P. and Leonida, L. (2009). 'Non-market effects of education on crime: evidence from Italian regions', *Economics of Education Review*, Vol. 28, pp. 11–17.

Cunha, F. and Heckman, J. J. (2007). 'The technology of skill formation', *American Economic Review*, Vol. 97, pp. 674–701.

Ehrlich, I. (1973). 'Participation in illegitimate activities: a theoretical and empirical investigation', *Journal of Political Economy*, Vol. 81, pp. 521–565.

Freeman, R. B. (1991). *Crime and the Employment of Disadvantaged Youths. No. 3875*. National Bureau of Economic Research.

Freeman, R. B. (1994). *Crime and the Job Market. No. 4910*. National Bureau of Economic Research.

Gaure, S., Røed, K., and Zhang, T. (2007). 'Time and causality: a Monte Carlo assessment of the timing-of-events approach', *Journal of Econometrics*, Vol. 141, pp. 1159–1195.

Grogger, J. (1998). 'Market wages and youth crime', *Journal of Labor Economics*, Vol. 16, pp. 756–791.

Hansen, K. (2003). 'Education and the crime-age profile', *The British Journal of Criminology*, Vol. 43, pp. 141–168.

Heckman, J. J. (2000). 'Policies to foster human capital', *Research in Economics*, Vol. 54, pp. 3–56.

Heckman, J. J. and Singer, B. H. (1984). 'A method for minimizing the impact of distributional assumptions in econometric models for duration data', *Econometrica*, Vol. 52, pp. 271–320.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). 'The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior', *Journal of Labor Economics*, Vol. 24, pp. 411–482.

Heller, S., Pollack, H. A., Ander, R. and Ludwig, J. (2013). *Preventing Youth Violence and Dropout: A Randomized Field Experiment. No. 19014*. National Bureau of Economic Research.

Hindelang, M. J., Hirschi, T. and Weis, J. G. (1981). *Measuring Delinquency*. Sage Publications, Beverly Hills, CA.

Hjalmarsson, R. (2008). Criminal justice involvement and high school completion. *Journal of Urban Economics*, 63, 613–630.

Hoffmann, J., Erickson, L., and Spence, K. (2013). 'Modeling the association between academic achievement and delinquency: an application of interactional theory', *Criminology*, Vol. 51, pp. 629–660.

Jacob, B. A. and Lefgren, L. (2003). 'Are idle hands the devil's workshop? Incapacitation, concentration, and juvenile crime', *American Economic Review*, Vol. 93, pp. 1560–1577.

Jensen, R. (2010). 'The (perceived) returns to education and the demand for schooling', *Quarterly Journal of Economics*, Vol. 125, pp. 515–548.

Johnson, A. W.(1997). *Mentoring at-Risk Youth: A Research Review and Evaluation of the Impacts of the Sponsor-a-Scholar Program on Student Performance*. University of Pennsylvania dissertation paper AAI9727243.

Kann, L., McManus, T., Harris, W. A., Shanklin, S. L., Flint, K. H., Queen, B., Lowry, R., Chyen, D., Whittle, L., Thornton, J., Lim, C., Bradford, D., Yamakawa, Y., Leon, M., Brener, N., and Ethier, K. A. (2018). 'Youth risk behavior surveillance - United States, 2017', *Morbidity and Mortality Weekly Report Surveillance Summaries*, Vol. 67, pp. 1–114.http://dx.doi.org/10.15585/mmwr.ss6708a1external icon.

Kemple, J. J. and Snipes, J. C. (2000). *Career Academies: Impacts on Students' Engagement and Performance in High School*. Manpower Demonstration Research Corporation, New York.

Kupchik, A. (2010). *Homeroom Security: School Discipline in an Age of Fear*. New York University Press, New York.

Levitt, S. D. and Lochner, L. (2001). 'The determinants of juvenile crime', in Gruber, J. (ed), *Risky Behavior among Youths: An Economic Analysis*, Chicago: University of Chicago Press, pp. 327–374..

Lochner, L. (2004). 'Education, work, and crime: a human capital approach', *International Economic Review*, Vol. 45, pp. 811–843.

Lochner, L. (2007). Education and Crime. Working Paper, University of Western Ontario.

Lochner, L. and Moretti, E. (2004). 'The effect of education on crime: evidence from prison inmates, arrests, and self-reports', *American Economic Review*, Vol. 94, pp. 155–189.

Loeber, R. (1982). 'The stability of antisocial and delinquent child behavior: a review', *Child Development*, Vol. 53, pp. 1431–1446.

Luallen, J. (2006). 'School's out... forever: a study of juvenile crime, at-risk youths and teacher strikes', *Journal of Urban Economics*, Vol. 59, pp. 75–103.

Machin, S., Marie, O., and Vujić, S. (2011). 'The crime reducing effect of education', *Economic Journal*, Vol. 121, pp. 463–484.

Machin, S., Marie, O., and Vujić, S. (2012). 'Youth crime and education expansion', *German Economic Review*, Vol. 13, pp. 366–384.

Machin, S. and Meghir, C. (2004). 'Crime and economic incentives', *Journal of Human Resources*, Vol. 39, pp. 958–979.

McCord, J., Spatz Widom, C., and Crowell, N. A. (2001). *The Development of Delinquency*. National Academy Press, Washington, DC.

McVicar, D., Moschion, J., and van Ours, J. C. (2019). 'Early illicit drug use and the age of onset of homelessness', *Journal of the Royal Statistical Society. Series A: Statistics in Society*, Vol. 182, pp. 345–372.

Merlo, A. and Wolpin, K. I. (2015). 'The transition from school to jail: youth crime and high school completion among black males', *European Economic Review*, Vol. 79, pp. 234–251.

Mocan, N. and Rees, D. I. (2005). 'Economic conditions, deterrence and juvenile crime: evidence from micro data', *American Law and Economics Review*, Vol. 7, pp. 319–349.

Moffitt, T. E. (1993). 'Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy', *Psychological Review*, Vol. 100, pp. 674–701.

Moschion, J. and van Ours, J. C. (2019). 'Do childhood experiences of parental separation lead to homelessness? *European Economic Review,* 111, 211–236.

Mroz, T. A. (1999). 'Discrete factor approximations in simultaneous equation models: estimating the impact of a dummy endogenous variable on a continuous outcome', *Journal of Econometrics*, Vol. 92, pp. 233–274.

Mroz, T. A., Picone, G., Sloan, F., and Yashkin, A. (2016). 'Screening for a chronic disease: a multiple stage duration model with partial observability', *International Economic Review*, Vol. 57, pp. 915–934.

Murnane, R. J., Willett, J. B., and Levy, F. (1995). 'The growing importance of cognitive skills in wage determination', *Review of Economics and Statistics*, Vol. 77, pp. 251–266.

Neal, D. A. and Johnson, W. R. (1996). 'The role of premarket factors in black-white wage differences', *Journal of Political Economy*, Vol. 104, pp. 869–895.

Neumark, D. and Joyce, M. (2001). 'Evaluating school-to-work programs using the new NLSY', *Journal of Human Resources*, Vol. 36, pp. 666–702.

Nordberg, M. and Røed, K. (2009). 'Economic incentives, business cycles, and long-term sickness absence', *Industrial Relations*, Vol. 48, pp. 203–230.

Oreopoulos, P. and Salvanes, K. G. (2011). 'Priceless: the nonpecuniary benefits of schooling', *Journal of Economic Perspectives*, Vol. 25, pp. 159–84.

Rud, I., van Klaveren, C., Groot, W., and Van den Brink, H. M. (2018). 'What drives the relationship between early criminal involvement and school dropout? *Journal of Quantitative Criminology,* 34, 139–166.

Segal, C. (2013). 'Misbehavior, education, and labor market outcomes', *Journal of the European Economic Association*, Vol. 11, pp. 743–779.

Siennick, S. and Staff, J. (2008). 'Explaining the educational deficits of delinquent youths', *Criminology*, Vol. 46, pp. 609–635.

Sweeten, G., Bushway, S. D., and Paternoster, R. (2009). 'Does dropping out of school mean dropping into delinquency? *Criminology,* 47, 47–91.

Traag, T., Marie, O., and Van der Velden, R. K. W. (2014). 'The relationship between school performance, delinquency and early school-leaving', in Weerman, F. and Bijleveld, C. (eds), *Criminal Behaviour from School to the Workplace (Untangling the Complex Relations Between Employment, Education and Crime)*, London, UK; Routledge/Taylor and Francis Group, pp. 44–60.

Ward, S. and Williams, J. (2015). 'Does juvenile delinquency reduce educational attainment? *Journal of Empirical Legal Studies,* 12, 716–756.

Webbink, D., Koning, P., Vujić, S., and Martin, N. G. (2013). 'Why are criminals less educated than non-criminals? Evidence from a cohort of young Australian twins', *The Journal of Law, Economics, and Organization*, Vol. 29, pp. 115–144.

Webbink, D., Vujić, S., Koning, P., and Martin, N. G. (2012). 'The effect of childhood conduct disorder on human capital', *Health economics*, Vol. 21, pp. 928–945.

Williams, J. and Sickles, R. C. (2002). 'An analysis of the crime as work model: evidence from the 1958 Philadelphia birth cohort study', *Journal of Human Resources*, Vol. 37, pp. 479–509.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix (online): Variable construction and data reliability