I Z A

Bad Behavior:
Delinquency, Arrest and Early School Leaving

Shannon Ward
Jenny Williams
Jan C. van Ours

Bad Behavior: Delinquency, Arrest and Early School Leaving

Shannon Ward
University of Melbourne

Jenny Williams
University of Melbourne
and IZA

Jan C. van Ours
CentER, Tilburg University,
University of Melbourne, CEPR and IZA

ABSTRACT

Bad Behavior:
Delinquency, Arrest and Early School Leaving[*]

In this paper we investigate the effects of delinquency and arrest on school leaving using information on males from the National Longitudinal Survey of Youth 1997. We use a multivariate mixed proportional hazard framework in order to account for common unobserved confounders and reverse causality. Our key finding is that delinquency as well as arrest leads to early school leaving. Further investigation reveals that the effect of delinquency is largely driven by income generating crimes, and the effect of both income generating crime and arrest are greater when onset occurs at younger ages. These findings are consistent with a criminal capital accumulation mechanism. On the basis of our sample, we show that taking into account the proportion of young men affected by delinquency and arrest, that the overall reduction in education due to delinquency is at least as large as the reduction due to arrest. This highlights the need for crime prevention efforts to extend beyond youth who come into contact with the justice system.

Corresponding author:

Jan C. van Ours
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
E-mail: vanours@uvt.nl

1 Introduction

In the United States, the net annual burden of crime is estimated to exceed \$1 trillion (Anderson, 1999). A large part of this sum is attributed to the cost of incarcerating close to 1.6 million individuals in state and federal prisons, and roughly 61,000 juvenile offenders in residential placement facilities (Carson, 2014; Office of Juvenile Justice and Delinquency Prevention, 2013). The burden of incarcerating such a large part of the population extends beyond the financial costs of imprisonment, and includes decreased material and psychological wellbeing of the families of those imprisoned as well as lost productivity (Anderson, 1999; Bulow, 2014; Schwartz-Soicher et al., 2011; Christian et al., 2006). In response to this significant social and economic issue, attention has increasingly turned to identifying the factors that lead to crime and incarceration. What has emerged as a consistent and robust predictor of these outcomes in the adult population is a low level of education. For example, around two thirds of adult inmates have not graduated high school (Harlow, 2003). Recent research demonstrates that this relationship is causal, with higher levels of education reducing the likelihood of crime and incarceration amongst adults (Lochner and Moretti, 2004; Oreopoulos and Salvanes, 2009; Merlo and Wolpin, 2009). Juxtaposing this literature is a body of research that documents a trajectory leading to adult crime, in which the starting point is typically antisocial and delinquent behavior in youth (Moffitt, 1993, 2006; Sampson and Laub, 2005; Williams and Sickles, 2002; Merlo and Wolpin, 2009). Taken together, these two bodies of research suggest that the relationship between delinquency, education and adult crime is a dynamic one, in which choices and outcomes regarding delinquency and schooling in youth are pivotal. Yet the nature of the relationship between delinquency and schooling remains unclear. Illuminating this issue is the focus of our research.

As discussed in the following section, several recent studies have made useful contributions to understanding the relationship between delinquency and schooling. For example, arrests and incarceration in youth are found to reduce educational attainment (Hjalmarsson, 2008; Webbink et al., 2013; Kirk and Sampson, 2013; Aizer and Doyle Jr, 2015), while school attendance decreases contemporaneous arrests, reported incidences of crime, and prosecutions for property crime (Jacob and Lefgren, 2003; Luallen, 2006; Berthelon and Kruger, 2011; Anderson, 2014). The criminal or juvenile justice focus of this literature has clear and important policy implications. However, interactions between youths and the justice system result from the decision of youths to engage in delinquency. This decision has the potential to impact on schooling outcomes, whether or not arrest or incarceration eventuates. For example, in a model of crime that accounts for human and criminal capital accumulation, experience in delinquency builds criminal capital stock which, in turn, lowers the returns to education relative to crime in adulthood

(Lochner, 2004). This makes education relatively less attractive, and an early school exit more likely, even if no arrest occurs. As a consequence, delinquency may lead to early school leaving for those who do not come to the attention of the criminal justice system, in addition to those who do.

A key contribution of our paper is that we provide empirical evidence on the role of delinquency, in addition to arrest, in determining early school leaving. Knowledge of whether delinquency leads to early school leaving, and the importance of its effect relative to that of arrest, is of direct policy relevance. It is particularly informative on whether and to what extent crime prevention efforts should extend beyond youth who come into contact with the justice system. A further contribution of this research is that we provide evidence on differential effects of delinquency and arrest on school leaving by type of delinquent behavior, and timing of initiation into delinquency and first arrest. The importance of these findings are twofold. First, they identify the profiles of delinquents who are at risk of early school leaving. Second, they provide important insights into mechanisms through which delinquency impacts the school leaving decision. Each of these aspects are critical for the development of policies that are effective at mitigating the effects of delinquent behavior, and for targeting them at those who will benefit the most.

There are two empirical challenges in determining the causal effects of delinquency and arrest on school leaving. First, unobserved characteristics such as time and risk preferences, and cognitive and non-cognitive abilities, that affect early school leaving are also likely to determine delinquency and arrest, rendering them endogenous to the school leaving decision. For example, Heckman et al. (2006) show that higher levels of noncognitive ability increase the probability that a person will graduate from high school and four-year college, while decreasing the probability of participating in crime. The second issue is that of reverse causality, with prior research showing that in addition to arrest leading to school leaving, school leaving increases the likelihood of arrest. In order to address these issues, we focus on the transitions into delinquency and arrest, and the transition out of school using a multivariate mixed proportional hazard (MMPH) framework. This approach is advantageous in that identification of treatment effects is achieved by exploiting differences in the timing of initiation into delinquency and first arrest relative to leaving school. Within this multivariate framework, we account for unobserved common confounders by introducing unobserved heterogeneity drawn from a joint distribution into the hazard rates for school leaving, the onset of delinquency, and first arrest. We account for reverse causality by modelling the relationships between the transitions into delinquency and out of school, and first arrest and school leaving as bi-directional.

We base our empirical analysis on data from the National Longitudinal Survey of Youth 1997 (NLSY97), a panel study of youths residing in the United States. These data are particularly

well suited to our purpose, containing retrospective in addition to annually collected information on delinquent behavior, arrests and educational attainment that allows us to discern the age at which initiation into delinquent activity occurs, the age at first arrest and the age at which an individual leaves school for the first time. This information is critical to our estimation and identification framework.

As with previous research, our analysis focuses on males. We find that after accounting for unobserved common confounders and reverse causality, both arrest and delinquency lead to early school leaving. We find the effect of arrest on the likelihood an individual leaves school is approximately twice the magnitude of the effect of delinquency. However, in the NLSY97 sample, there are twice as many males enrolled in school at age 17 that have been delinquent and not arrested compared to the number who are enrolled in school and have been arrested. As a consequence, the overall effect of delinquency on educational attainment is at least as large as that of arrest. Further analysis reveals that the effect of delinquency on the school leaving rate is largely driven by initiation into delinquency that is income generating, and that earlier initiation has a larger impact on school leaving than later initiation. These findings are consistent with predictions from a capital accumulation theory of crime. In contrast, initiation into non-income generating delinquency only has an effect on school leaving if initiation occurs at age 18 or older, suggesting that there are salient or vulnerable ages with regard to the school leaving decision. Our results with respect to arrest are consistent with previous studies. We find that first arrest prior to age 18 increases the transition rate out of school, with the magnitude of this effect being larger for those whose first arrest occurs at younger ages. Also consistent with previous studies, we find evidence of reverse causality in the relationship between school leaving and arrest, with early school leaving increasing the likelihood of arrest. We find no evidence, however, that early school leaving leads to initiation into delinquency.

The rest of the paper is laid out as follows. Section 2 offers background information, providing a discussion of the existing economics literature on criminal activity and school leaving. This is followed by Section 3, which introduces the National Longitudinal Survey of Youth 1997, provides descriptive statistics and discusses the data's key features. Section 4 presents the empirical model and estimation strategy used to identify the causal effects of delinquency and arrest on school leaving, the results of which are reported in Section 5. Section 6 follows with sensitivity analyses that include an investigation into the potential for reverse causality, along with several other more detailed analyses. Section 7 explores capital accumulation as a mechanism underlying the causal effects we identify. Section 8 concludes with a discussion of our findings and associated policy implications.

2 Literature Review

2.1 Theory

In the standard economic model of crime proposed by Becker (1968) and extended by Ehrlich (1973), individuals allocate resources to engaging in crime if and only 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 labor market earnings, essentially making crime a risky occupation.[1] Of course, returns to legitimate and illegitimate activities depend on the relevant capital stocks that have been accumulated in previous periods. In order to capture this dynamic aspect of criminal choice, Lochner (2004) develops a life cycle model of crime in which individuals allocate time to investing in human capital, work and crime so as to maximize expected lifetime earnings (from crime and work). While the focus of his model is on showing that post-school investment in human capital accumulation plausibly explains different age patterns in desistance for low and high skill crimes, Lochner (2004) also considers criminal capital accumulation that occurs via criminal experience.

As discussed by Lochner (2004), engaging in delinquency can affect school leaving in a capital accumulation framework because it builds criminal capital, which increases the expected future returns to crime relative to education. Early drop out is then the consequence of falling expected relative returns to education.[2] An additional implication of this framework 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. A further implication of the capital accumulation model of crime is that early initiation into income generating delinquency is predicted to have a larger effect on school leaving than late initiation, as those who initiate income-generating delinquency at younger ages have more periods over which to reap the rewards of their criminal capital stock.

There are several other mechanisms through which delinquency may affect educational attainment. For example, delinquency may affect schooling through the time constraint if time spent in delinquency reduces time spent studying. In addition, if delinquent behavior 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 may be perceived

---

[1]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.

[2]This is particularly relevant if the wages offered to less educated young men are relatively low, allowing crime to become a more lucrative alternative (Mustard, 2010; Machin and Meghir, 2004; Lochner, 2007; Freeman, 1994, 1991).

as harassment (Hjalmarsson, 2008; Kupchik, 2010; Segal, 2013). Less time spent studying or attending classes, and an increased level of animosity due to perceived differential treatment by teachers is likely to result in lower academic achievement and an early exit from school. There are, thus, several mechanisms via which engaging in delinquency may impact school leaving, over and above the effect of being arrested.

2.2 Empirical Literature

The literature investigating the impact of criminal activity in youth on educational attainment is both recent and small.[3] The focus of this literature is on the impact of arrest and incarceration, rather than on the impact of delinquency, on schooling. For example, Hjalmarsson (2008) finds that arrest and incarceration by age 16 decrease the probability of high school graduation by age 19, although only the effect of incarceration is found to be robust. Using data on twins, Webbink et al. (2013) report that early arrests decrease educational attainment and lower the probability of completing high school for fraternal twins. Kirk and Sampson (2013) find that youth who are arrested are more likely to drop out of high school and less likely to enrol in college. Aizer and Doyle Jr (2015) find that juvenile incarceration decreases high school graduation. The overall conclusion to be drawn from this literature is that interactions with the juvenile justice system have a detrimental effect on educational attainment in general, and graduation from high school in particular.

There are also studies showing that the relationship between education and interactions with the juvenile justice system runs in the opposite direction. These studies investigate the contemporaneous effect of attending school on youths arrests, reported crime incidences and prosecutions, finding school attendance reduces property crime. For example, Jacob and Lefgren (2003) find that an additional day of school attendance reduces the reported incidences of juvenile property crime, but increases juvenile violent crime. Similarly, Luallen (2006) finds that an additional school day decreases juvenile arrests for property crime and increases juvenile arrests for violent crime, with a net effect of increased total juvenile crime. Berthelon and Kruger (2011) investigate the effect of extending the length of the school day (through the gradual, nationwide implementation of a full-day school reform) on contemporaneous municipal juvenile prosecution rates, finding that the reform decreased total, property and violent crime. Anderson (2014) finds increasing the minimum dropout age from 16 to 18 decreases property, violent and overall arrest rates for those aged 16 to 18. Overall, this evidence suggests that schooling reduces youthful property crime. The evidence is somewhat mixed with respect to violent crime.

---

[3]A more substantial literature investigates the impact of educational attainment on crime, arrest and incarceration in adulthood (Lochner and Moretti, 2004; Oreopoulos and Salvanes, 2009; Merlo and Wolpin, 2009; Buonanno and Leonida, 2009; Machin et al., 2011, 2012).

This literature has produced important empirical findings that have useful policy implications. Nonetheless, the picture they provide is partial and incomplete. This is because, in focusing on the relationship between schooling and arrests or incarceration, these studies ignore the fact that interactions with the juvenile justice system arise as a result of an individual's decision to offend. The importance of this decision is highlighted by Lochner's model, which illustrates how individuals who choose to allocate time to crime potentially experience long-lived consequences both in terms of their offending behavior and educational attainment, even in the absence of arrest and incarceration. This brings us to the contribution of the current paper. In this research, we seek to answer the question of whether, and to what extent, engaging in delinquent behavior causally impacts school leaving amongst those who are not apprehended, in addition to those who are.

In order to address this question we focus on the first time an individual is delinquent, is arrested, and leaves school, and develop an empirical strategy (discussed in detail in Section 4) based on a multivariate hazard approach in which unobserved heterogeneities are drawn from a joint distribution. Strategies employed by previous studies to identify the effects of arrest and incarceration on schooling include fixed effects, propensity score matching and an instrumental variable approach.[4] While a fixed effects approach controls for time invariant unobserved characteristics, it does not account for reverse causality, which previous studies show to be an issue. Propensity score matching does not allow for reverse causality, and relies on a difficult to justify assumption of conditional independence. An instrumental variable approach requires variables that affect the probability of delinquency and arrest, but not the probability of leaving school. However, valid instruments are unlikely to exist for our particular case.[5] In contrast, our approach is innovative in that we take advantage of the timing of the initiation of delinquency and first arrest relative to school leaving to disentangle the direction of causality in the relationship between the onset of delinquency and school leaving, and that between first arrest and school leaving. Our framework permits us to account for both potential unobserved confounders and reverse causality in determining the effect of delinquency on school leaving, over and above that of arrest. It does, however, require very detailed information on the timing of these outcomes, which the National Longitudinal Survey of Youth 1997 provides.

---

[4]Aizer and Doyle Jr (2015) use an instrumental variable strategy that relies on exogenous variation in juvenile incarceration resulting from the random assignment of juveniles to judges with different tendencies to sentence youths to imprisonment.

[5]Note that as expenditure on criminal justice and education is jointly determined in state budget processes, criminal justice policy variables are unlikely to be conditionally independent of education outcomes and hence, unsuitable to serve as instruments.

3 Data and Measures

3.1 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 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 and black people (N=2,236). We use both subsamples for our analysis, and perform a sensitivity analysis using only the cross-sectional sample. As with previous research, our analysis focuses on males as they commit the majority of delinquent acts (Grogger, 1998; Williams and Sickles, 2002). Of the 4,599 males interviewed for the NLSY97, we exclude 111 who had missing information on key variables.[6] This results in a final sample of 4,488 males.

3.2 Outcomes

The outcomes of interest for this study are the age at which an individual initiates delinquent behavior, the age at which an individual is first arrested, and the age at which an individual first leaves school (defined as primary, secondary or tertiary education). In this section, we briefly describe each of our outcome variables. A more detailed description of their construction is provided in the Appendix.

The NLSY97 elicits information on the following delinquent acts: whether the individual has stolen more than $50 worth, committed other property crimes, sold drugs, destroyed property, or attacked someone. In wave 1, respondents were asked the age at which they first committed each of the enumerated delinquent acts. We construct the age at which an individual initiates delinquent behavior by combining the wave 1 information with responses from each subsequent wave as to whether they had committed each act since the last interview. We also use this information to construct the ages at which an individual commits his first income generating delinquent act (which includes stealing more than $50 worth, other property crimes, and selling drugs) and his first non-income generating act (defined as destroying property or attacking others).

---

[6]Of the 111 individuals, 48 were dropped because they have delinquency processes that are right censored before reporting being arrested. This largely occurs because from round 8 onwards, the delinquency questions are only asked of respondents who had ever reported being arrested and a control group of approximately 10% of the respondents for comparison, whereas the questions on arrest are asked throughout all 13 waves.

Similarly, wave 1 of the NLSY97 asks individuals the age at which they were first arrested, and subsequent waves enquire as to whether they have been arrested since the last interview. This information is combined to construct the age at which an individual is first arrested. Note that the survey does not ask respondents the offence for which they are arrested, simply whether they are arrested. For this reason we do not distinguish between arrest for income or non-income generating acts.

We construct the age at which a respondent first leaves school (defined as primary, 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 enrolment status at the time of each interview. If the individual leaves school 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.

3.3   Explanatory Variables

Previous research has found that the individual-level determinants of crime and educational choice include gender, age, race, religion, cognitive ability, local environment, family background and structure, and parental education (Mocan and Rees, 2005; Levitt and Lochner, 2001). The NLSY97 collects rich information on the respondents' characteristics and their family backgrounds such that we are able to control for these important determinants.[7] Specifically, our set of controls includes 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 experiencing puberty before the age of 12; an indicator for attending a private or parochial school (attending a public school is the omitted category); and a set of indicators for year of birth. We control for the following parental characteristics: mother was a teen at the respondent's birth; the responding parent (to the parent questionnaire administered in round 1) is very religious; and parents' education (of the mother and father separately). We also control for the following variables related to family structure and household characteristics: number of siblings (younger and older separately); and parents present in household (separate indicators for mother and father). In addition, we 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). Detailed information on the construction of the controls is provided in

---

[7]However, we are unable to account for county-level characteristics (such as the unemployment rate and arrest rates) as the NLSY97 public-use files do not include geographic variables such as state or county.

the Appendix. Sample means, minima and maxima are presented in Table 1.

3.4 Reliability of Data

The NLSY97 are self-reported data and this raises several issues around its reliability. Previous research has evaluated the reliability of self-reported crime data, finding that self-reported data and official crime data generally yield similar information (Mocan and Rees, 2005; Elliott and Voss, 1974; Hindelang et al., 1981; Farrington, 1973). With regard to the data used in this study, Lochner (2007) compares the NLSY97 data on arrest rates (defined as the total number of self-reported arrests divided by the number of self-reported crimes) with data from the Federal Bureau of Investigation's *Uniform Crime Reports*, and finds the NLSY97 self-reported crime data to be generally consistent with official statistics, after adjusting for non-reporting to the police by victims. Moreover, as noted by Levitt and Lochner (2001), given that many crimes are not reported to the police and that most reported crimes do not lead to an arrest, criminal participation is likely to be much more prevalent than naively implied by official arrest data. Self-reported crime data are, therefore, considered both useful and informative.

An additional issue is that of recall error. This is particularly salient given the outcome variables used in our analyses are the age at which a respondent first engages in delinquency, the age he is first arrested and the age he first leaves school. Studies based on cross-sectional data rely on respondents' memories of when (and the order in which) these events occurred, and are likely to be subject to recall error. In contrast, the NLSY97 data are longitudinal, and the period between waves is relatively short (one year). As a consequence, the risk of recall error is substantially reduced in these data.[8] This allows greater accuracy regarding the ordering and timing of events under analysis.

Lastly, there may be concerns about differential reporting of delinquency and arrest by race. Hindelang et al. (1981) suggest that young men who underreport criminal or delinquent behavior are more likely to be black than white. Lochner and Moretti (2004) compare the self-reported crime rate in the NLSY79 with incarceration data, inferring that black dropouts rather than black high school graduates or whites tend to underreport their criminal behavior. Lochner and Moretti (2004) deal with this issue by dividing their sample by race and estimating results separately for blacks and whites. On the other hand, Merlo and Wolpin (2009) use self-reported data on crime, schooling and work from the NLSY97 and base their analysis on black males only. Given the potential for differential reporting, we investigate this issue in a sensitivity analysis in Section 6.3.

---

[8]Recall may remain an issue for those who first made transitions before wave 1 when respondents were 12-18 years old.

3.5 Descriptive Statistics

Table 1 provides descriptive statistics for our data. It shows that 68% of the males in our sample have engaged in some type of delinquent activity, while 32% have not. Similarly, and keeping in mind that individual can engage in both income and non-income generating delinquency, 43% of our sample have committed at least one income generating delinquent act, and 62% have committed at least one non-income generating delinquent act. Lastly, we find that 44% of our sample have been arrested and 94% left school during the observation period. Table 1 also provides evidence on the timing of events, showing the average age of initiating delinquency to be 12 years old and much younger than the average age of first arrest at age 17, which in turn occurs before the typical school leaving age of 19.

Figure 1 provides more detailed information on transitions into delinquency and arrest, and out of school.[9] Specifically, it shows the hazard rates for the transition out of school and the transitions into first arrest and initiation into delinquency. Hazard rates for initiation into delinquency are shown for any type of delinquency as well as for income generating and non-income generating delinquency separately. The figure shows that the hazard rates for any delinquency and non-income generating delinquency peak at age 14, the hazard rate for income generating delinquency peaks slightly later at age 15, and the hazard rate for first arrest peaks later still at age 18.[10] 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 summary, the hazard rates for the transitions into (income, non-income and any type of) delinquent behavior peak at ages younger than the peak in the hazard for arrest, which in turn occurs before the first peak in the hazard for school leaving. This suggests that the typical order of events is initiation into delinquent behavior, arrest, followed by school leaving.

The cumulative probability distributions for first arrest, the initiation of delinquency and school leaving are presented in Figure 2, with distributions once again presented separately for income generating delinquency, non-income generating delinquency, and any type of delinquency. The figure shows that the probability of having engaged in income (non-income) generating

---

[9]Distributions for the ages at which individuals first leave school, are first arrested and initiate delinquency (with separate distributions for any, income and non-income generating activity) are presented in Appendix Table A.1.

[10]At age 7, roughly 8% of respondents report having engaged in delinquent behavior. 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 delinquency and 188 for non-income generating delinquency (given that income and non-income generating are not mutually exclusive categories).

delinquent activity by age 15 is 0.25 (0.48), increasing to 0.38 (0.58) by age 18, 0.43 (0.61) by age 20, and to 0.46 (0.64) by age 25. Importantly, the figure illustrates that respondents appear unlikely to take up delinquent behavior of any kind after age 19. This suggests that there are a minimum of two types of individual with regard to delinquency: those who are susceptible to delinquent behavior and those who are not. Similarly, the cumulative probability distribution for first arrest shows that the probability of having been arrested by age 15 is 0.15, increasing to 0.29 by age 18, 0.35 by age 20, and to 0.44 by age 25. From the figure, we can see that respondents are unlikely to be arrested for the first time after age 25, suggesting two types of individual with regard to arrest: those susceptible and those not. While the cumulative probability distributions for first arrest and the initiation of delinquency reflect that individuals may or may not ultimately initiate delinquency or be arrested, the cumulative probability distribution for school leaving shows that eventually everyone will leave school. Specifically, it shows that the probability of having left school by age 18 is 0.40. This increases to 0.62 by age 19, 0.71 by age 20, 0.83 by age 22, and to 0.97 by age 25. The distribution increases steeply over ages 17, 18 and 19, illustrating that this is when most students leave school.

Further evidence on the timing of first delinquency and arrest relative to school leaving is provided by Table 2. Column (a) presents the probabilities associated with the order in which the initiation of delinquency and school leaving occur. Column (b) does the same for the order of initiation of income generating delinquency and school leaving, column (c) for the initiation of non-income generating delinquency and school leaving, and column (d) for first arrest and school leaving. Column (a) shows that 63% of the sample initiate delinquency before leaving school, compared with only 3% who leave school prior to initiating delinquency, while 2% of our sample initiate delinquency and leave school at the same age. Similarly, in column (b) 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, while 3% do both at the same age. Column (c) 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, while 2% both leave school and initiate non-income delinquency at the same age. Lastly, column (d) 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, while 4% of the sample are arrested (for the first time) and leave school at the same age. Overall, the table provides evidence that the initiation of delinquency and first arrest typically precede school leaving, and that should a causal relationship exist, it would likely run from delinquency and arrest to school leaving.

4 Estimation

4.1 Empirical Framework

The primary objective of our empirical analysis is to determine the separate causal effects of delinquency and arrest on school leaving. As previously discussed, in order to do so we must account for unobserved common confounders that determine the transitions into delinquency, arrest and school leaving, and the potential for reverse causality in the relationships between delinquency and school leaving, and arrest and school leaving. Our strategy is based on a multivariate mixed proportional hazard (MMPH) framework. This entails modelling a system, 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 impact the school leaving transition rate. We account for unobserved common factors by allowing the unobserved heterogeneity terms entering each of the transition rates to be drawn from a joint distribution. We address the issue of reverse causality as part of a sensitivity analysis (in Section 6.1), in which we additionally permit school leaving to affect the transitions into delinquency and arrest.

A proof of the identification of the treatment effect in the multivariate mixed proportional hazard model is provided by Abbring and van den Berg (2003). They show that identification is achieved through the order in which events occur; for example, the timing of first arrest relative to school leaving. They also show that identification can be achieved without restrictions that are necessary in a propensity score matching or instrumental variable approach, such as conditional independence or the existence of valid instruments. This provides a considerable advantage to using the multivariate mixed proportional hazard approach.

Our empirical investigation begins by providing a point of comparison with previous literature by examining the impact of arrest on early school leaving. To do so, we use a bivariate mixed proportional hazard framework. We then build up to the baseline multivariate mixed proportional hazard model, in which transitions into delinquency in addition to arrest are modelled along with school leaving.

4.1.1 The bivariate mixed proportional hazard model

The hazard (or transition) rate for leaving school at age t conditional on observable characteristics x, the age at which an individual is first arrested t_a, and unobservable characteristics v, is specified as

$$\theta_s(t|x, t_a, v) \quad = \quad \lambda_s(t) \exp(x'\beta_s + \delta_a I(t_a < t) + v) \tag{1}$$

where the indicator function $I(t_a < t)$ equals one if the individual experienced his first arrest before age t. Individual duration dependence is represented by $\lambda_s(t)$, while β_s is a vector of parameters to be estimated. Note that arrest is only permitted to impact on school leaving at age t if it occurs at an age prior to age t. This is because we are unable to tell which event happened first if school leaving and arrest occur at the same age.

The key parameter of interest is the coefficient on prior arrest, δ_a. If it is positive ($\delta_a > 0$), then arrest increases transitions out of schooling, leading to lower educational attainment. If it is negative ($\delta_a < 0$), then arrest delays school leaving, leading to higher educational attainment. If δ_a is zero, then arrest has no effect on school leaving or educational attainment. The unobserved heterogeneity term, v, accounts for differences in propensities for school leaving.

Duration dependence is modelled flexibly using a step function, and is specified as

$$\lambda_s(t) = \exp\left(\sum_{k=1}^{12} \lambda_{s,k} I_k(t)\right) \tag{2}$$

where there are twelve duration categories ($k = 1, ..., 12$), eleven of which represent specific ages (14, ..., 24), while the last interval is for ages greater than 24. $I_k(t)$ 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.

The conditional density function for the completed durations until the individual leaves school is specified as

$$f_s(t|x, t_a, v) = \theta_s(t|x, t_a, v) \exp\left(-\int_0^t \theta_s(z|x, t_a, v)dz\right) \tag{3}$$

We assume that those who have not left school during the observation period have a right-censored duration until school leaving.

We similarly specify the hazard rate for first arrest at age t, conditional on observable characteristics x, and unobservable characteristics ε as,

$$\theta_a(t|x, \varepsilon) = \lambda_a(t) \exp(x'\beta_a + \varepsilon) \tag{4}$$

where β_a is a vector of parameters to be estimated. We assume that individuals are at risk of being arrested from age ten, and model individual duration dependence using the step function,

$$\lambda_a(t) = \exp\left(\sum_{j=1}^{16} \lambda_{a,m} I_m(t)\right) \tag{5}$$

where there are sixteen duration categories ($m = 1, ..., 16$), fifteen of which represent specific ages (10, ..., 24), while the last interval is for ages greater than 24. $I_m(t)$ are a set of m dummy variables equal to one if the individual is arrested for the first time in duration category m, and zero otherwise. We set $\lambda_{a,1} = 0$, given that we estimate a constant term in $x'\beta_a$.

The conditional density function for the completed durations until an individual's first arrest is specified as

$$f_a(t|x, \varepsilon) = \theta_a(t|x, \varepsilon) \exp\left(-\int_0^t \theta_a(z|x, \varepsilon)dz\right) \tag{6}$$

We assume that those who have not been arrested during the period under observation have a right-censored duration until arrest.

The unobserved component in the hazard rate for first arrest (ε) is potentially correlated with the unobserved component in the hazard rate for school leaving (v). To account for this, we specify (v, ε) as drawn from a discrete joint distribution function, $G(v, \varepsilon)$. In practise, we find two mass points in the distribution of unobserved heterogeneity entering the school leaving equation, and two mass points in the distribution of unobserved heterogeneity affecting the transition rate into first arrest. Therefore, the joint distribution, $G(v, \varepsilon)$ has up to four points of support (v_1, ε_1), (v_1, ε_2), (v_2, ε_1), (v_2, ε_2).[11] The four points of support, which may considered a distinct "types" of individuals, occur with the following probabilities:

$$\Pr(v = v_1, \varepsilon = \varepsilon_1) = p_1, \quad \Pr(v = v_1, \varepsilon = \varepsilon_2) = p_2$$
$$\Pr(v = v_2, \varepsilon = \varepsilon_1) = p_3, \quad \Pr(v = v_2, \varepsilon = \varepsilon_2) = p_4$$

with $0 \leq p_j \leq 1$ for $j = 1, ..., 4$, and $p_4 = 1 - p_1 - p_2 - p_3$.[12] We model these probabilities using a multinomial logit specification

$$p_j = \frac{\exp(\alpha_j)}{\sum_j \exp(\alpha_j)}$$

for $j = 1, 2, 3, 4$

Integrating out the unobserved heterogeneity leads to the following joint density function for the duration of time until first arrest t_a, and the duration of time until school leaving t_s,

---

[11] These result from concatenating the two masspoints in each of the school leaving and arrest equations.

[12] We model the second mass point (v_2, ε_2) differentially, such that

$$v_1 + v_{22} = v_2, \quad \varepsilon_1 + \varepsilon_{22} = \varepsilon_2$$

where v_1, ε_1 and v_{22}, ε_{22} are estimated.

conditional on x:

$$h_{s,a}(t_s, t_a|x) = \int_v \int_\varepsilon f_s(t|x, t_a, v) f_a(t|x, \varepsilon) dG(v, \varepsilon) \tag{7}$$

4.1.2 The multivariate mixed proportional hazard model

The bivariate mixed proportional hazard model presented above focuses on the impact of arrest on early school leaving. Building on this, we now outline our baseline multivariate mixed proportional hazard model, which allows delinquency, in addition to arrest, to impact on school leaving. In order to account for the impact of initiation into delinquency as well as that of first arrest, the hazard rate for school leaving (originally equation (1)) is augmented as follows

$$\theta_s(t|x, t_d, t_a, v) = \lambda_s(t) \exp(x'\beta_s + \delta_d I(t_d < t) + \delta_a I(t_a < t) + v) \tag{8}$$

where t_d is the age at which an individual first engages in delinquency and the indicator function $I(t_d < t)$ equals one if the individual initiated delinquency before age t. Note that, as with arrest, only delinquency that occurs at an age prior to t is permitted to affect school leaving at age t. The parameter δ_d, indicates whether prior delinquency has a positive ($\delta_d > 0$), negative ($\delta_d < 0$) or no effect ($\delta_d = 0$) on school leaving.

Accounting for the effect of delinquency, in addition to arrest, the conditional density function for the completed durations until the individual leaves school (originally equation (3)) is

$$f_s(t|x, t_d, t_a, v) = \theta_s(t|x, t_d, t_a, v) \exp\left(-\int_0^t \theta_s(z|x, t_d, t_a, v) dz\right) \tag{9}$$

We add to the system of equations the hazard rate for delinquency. Conditional on observable characteristics x, and unobservable characteristics u, the hazard for initiating into delinquency at age t is given by:

$$\theta_d(t|x, u) = \lambda_d(t) \exp(x'\beta_d + u) \tag{10}$$

where β_d is a vector of parameters to be estimated. We assume that individuals are at risk of engaging in delinquent behavior from age seven, which is around the age at which they start school,[13] and model individual duration dependence using the step function,

$$\lambda_d(t) = \exp\left(\sum_{j=1}^{19} \lambda_{d,j} I_j(t)\right) \tag{11}$$

---

[13]Depending on the state, schooling in the US is compulsory over an age range that starts between five and eight and ends between sixteen and eighteen.

where there are nineteen duration categories ($j = 1, ..., 19$), eighteen of which represent specific ages (7, ..., 24), while the last interval is for ages greater than 24. $I_j(t)$ are a set of j dummy variables equal to one if the individual initiates delinquent activity in duration category j, and zero otherwise. Again, we set $\lambda_{d,1} = 0$ as we estimate a constant term in $x'\beta_d$.

The conditional density function for the completed durations until the onset of delinquent behavior is specified as

$$f_d(t|x, u) = \theta_d(t|x, u) \exp\left(-\int_0^t \theta_d(z|x, u)dz\right) \tag{12}$$

We assume that those who have not initiated delinquency during the period under observation have a right-censored duration until the onset of delinquency.

The hazard rate for first arrest and the conditional density function for completed durations until first arrest are as previously specified (equations (4) and (6), respectively). To account for the potential correlation in unobserved heterogeneity determining school leaving, delinquency and arrest, we specify that (u, v, ε) are drawn from a discrete joint distribution, $W(u, v, \varepsilon)$. 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. This implies that the joint distribution of unobserved heterogeneity has up to eight points of support, and with the following probabilities:

$$\Pr(u = u_1, v = v_1, \varepsilon = \varepsilon_1) = p_1, \quad \Pr(u = u_1, v = v_1, \varepsilon = \varepsilon_2) = p_2$$
$$\Pr(u = u_1, v = v_2, \varepsilon = \varepsilon_1) = p_3, \quad \Pr(u = u_1, v = v_2, \varepsilon = \varepsilon_2) = p_4$$
$$\Pr(u = u_2, v = v_1, \varepsilon = \varepsilon_1) = p_5, \quad \Pr(u = u_2, v = v_1, \varepsilon = \varepsilon_2) = p_6$$
$$\Pr(u = u_2, v = v_2, \varepsilon = \varepsilon_1) = p_7, \quad \Pr(u = u_2, v = v_2, \varepsilon = \varepsilon_2) = p_8$$

noting that $0 \leq p_j \leq 1$ for $j = 1, ..., 8$, and $p_8 = 1 - p_1 - p_2 - p_3 - p_4 - p_5 - p_6 - p_7$.[14] We, again, model these probabilities using a multinomial logit specification

$$p_j = \frac{\exp(\alpha_j)}{\sum_j \exp(\alpha_j)}$$

for $j = 1, ..., 8$

---

[14]We model the second mass point $(u_2, v_2, \varepsilon_2)$ for each behavior differentially, such that

$$u_1 + u_{22} = u_2, \quad v_1 + v_{22} = v_2, \quad \varepsilon_1 + \varepsilon_{22} = \varepsilon_2$$

where u_1, v_1, ε_1 and $u_{22}, v_{22}, \varepsilon_{22}$ are estimated.

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

$$h_{d,s,a}(t_a, t_s, t_d|x) = \int_u \int_v \int_\varepsilon f_s(t|x, t_d, t_a, v) f_a(t|x, \varepsilon) f_d(t|x, u) dW(\varepsilon, v, u) \tag{13}$$

5 Empirical Results

5.1 Preliminary Analysis

Past studies investigating the relationship between delinquency and schooling in youth have focused on the impact of arrest on schooling, finding that arrest reduces educational attainment (Hjalmarsson, 2008; Webbink et al., 2013; Kirk and Sampson, 2013). In order to facilitate a comparison between our study and the literature, we provide a preliminary analysis that addresses this same question.

Table 3 presents the maximum likelihood estimates of a bivariate mixed proportional hazard model for school leaving and first arrest, in which arrest impacts on school leaving, and the unobserved heterogeneity in each process is potentially correlated. The first column provides estimates for the arrest equation, and the second column for the school leaving equation. The coefficient on arrest in the equation for school leaving is the parameter of interest.[15] Panel (a) of Table 3 reports the parameter estimates assuming that the unobserved heterogeneity terms in the arrest and school leaving processes are correlated, panel (b) reports key estimates assuming the unobserved heterogeneities are independent, and panel (c) reports the joint distribution of unobserved heterogeneity implied by the estimates in panel (a).

Starting with panel (a), the estimates suggest that arrest increases the rate at which young men leave school by 52% (100(exp(0.42)-1)), a result consistent with previous studies showing that arrest reduces educational attainment. The results also show that the effects of the explanatory variables are generally as expected. In terms of the transition into first arrest, the likelihood is lower for those who have higher ability as measured by the adjusted CAT-ASVAB score, those who attended private or parochial schools (with public schools as the comparison

---

[15]Our set of controls includes: 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 experiencing puberty before the age of 12; an indicator for attending a private or parochial school; a set of indicators for year of birth; mother was a teen at the respondent's birth; responding parent is very religious; parents' education; number of siblings (younger and older separately); parents present in household (separate indicators for mother and father); 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).

category), those who are Hispanic (relative to those who are non-black and non-Hispanic), those with a very religious parent who responded to the parent questionnaire, those who have a father with at least a high school education and a mother who has graduated high school but not college (compared with having a mother or a father who has not completed high school), and those whose mothers and fathers are present in the household. The likelihood of arrest is higher for those who enter puberty early, those whose mothers were teenagers at the respondents' births, those who live in the suburbs or city (compared to those who live in a rural area), and for those who live in the North (compared with those who live in the Northeast).

In terms of school leaving, a lower, and therefore later, transition out of school is associated with those who have higher ability as measured by the adjusted CAT-ASVAB score, those who attended private or parochial schools (with public schools as the comparison category), those who are black (relative to those who are non-black and non-Hispanic), those whose responding parent (to the parent questionnaire) is very religious, those with a mother or father with at least a high school education (compared with having a mother or father who has not completed high school), those whose mother and father are present in the household, and those who reside in the suburbs (compared to those who live in a rural area). The school leaving rate is higher for those whose mothers were teens at the respondents' births, for those with a greater number of (younger and older) siblings, and for those who reside in the South (with Northeast being the comparison category).

With respect to unobserved heterogeneity, we identify two types of individuals in the arrest equation and two types in the school leaving equation. For the arrest equation, the first and second mass points are estimated to be -3.34 and -5.17 ($= -3.34 - 1.83$), respectively. This indicates that the first type (with a mass point of -3.34) is more susceptible to arrest than the second type (with mass point -5.17). The first and second mass points for the school leaving equation are estimated to be -5.39 and -7.15 ($= -5.39 - 1.76$), representing individuals with (relatively) higher and lower susceptibilities to school leaving, respectively. Given two types for both school leaving and arrest, the joint distribution of unobserved heterogeneity has between two (in the case of perfect correlation) and four potential points of support.

As shown in panel (c) of Table 3, we identify three points of support in the joint distribution of unobserved heterogeneity. The table shows the percentage of the sample that represents each of the three types of individual. The first type is represented by 51% of the sample and is highly susceptible to both arrest and early school leaving. The second type has little susceptibility to arrest and a high susceptibility to early school leaving, and represents 39% of the sample. Lastly, 10% of the sample has a low susceptibility to both arrest and early school leaving.

Panel (b) of Table 3 presents results for a bivariate hazard model in which the unobserved

19

components of the school leaving and arrest transitions are assumed to be independent. Using an LR test to compare the specifications in panels (a) and (b) confirms that the unobserved heterogeneities determining school leaving and first arrest are not independent, with a test statistic of 23.2 exceeding the critical value of 3.8 for a χ^2_1 distribution at the 5% level of significance. This is indicated in the table by the stars (**) following the value of the LR test statistic. Comparing the correlated model (in panel (a)) with the independent model (in panel (b)) of Table 3 shows that ignoring correlation between the unobserved heterogeneity terms results in an upwardly biased coefficient estimate of the effect of arrest on school leaving. Specifically, failing to account for the endogeneity of arrest results in an overestimate of its effect on school leaving of 82% (100(exp(0.60)-1)), compared with the estimate from the correlated model of 52% (100(exp(0.42)-1)). The upwardly biased estimate produced by the independent hazard model suggests a positive correlation in unobserved heterogeneity. This means that the unobserved characteristics that increase the likelihood of arrest also increase the likelihood of early school leaving.

5.2 Delinquency, Arrest and School Leaving

Consistent with previous studies, we find that arrest leads to an early exit from school. Having established this point of comparison with the literature we now address the question of whether, after accounting for the impact of arrest, delinquency itself has an effect on school leaving. In order to do so, we estimate a multivariate system of equations for initiation into delinquency, first arrest and school leaving, in which delinquency and arrest separately impact school leaving, and the unobserved heterogeneities impacting the three processes are permitted to be correlated – a particularly appropriate approach given the high likelihood of correlation between unobservables determining delinquency and arrest.

The results from the multivariate hazard model are presented in Table 4. The first column provides estimates for the delinquency equation, the second column for the arrest equation, and the third column for the school leaving equation. Following the format of Table 3, panel (a) of Table 4 reports the parameter estimates assuming that the unobserved heterogeneity terms in the delinquency, arrest and school leaving processes are correlated, panel (b) reports key estimates assuming the unobserved heterogeneities are independent, and panel (c) reports the joint distribution of unobserved heterogeneity implied by the estimates in panel (a).

The parameters of interest are the coefficients on delinquency and arrest in the equation for school leaving. As can be seen from the third column of Table 4, we find that arrest increases the school leaving rate by 57% (100(exp(0.45)-1)) and delinquency increases it by 31% (100(exp(0.27)-1)). These effects are statistically significant at the 1% level. Comparing

the multivariate hazard model of delinquency, arrest and school leaving (in Table 4) with the bivariate hazard model between arrest and school leaving (in Table 3) shows that ignoring the delinquency process results in an underestimate of the effect of arrest on school leaving ($52\% = 100(\exp(0.42)-1)$), compared to when delinquency is accounted for ($57\% = 100(\exp(0.45)-1)$).

In terms of unobserved heterogeneity, panel (a) shows that we identify two types in each of the transitions out of school, into delinquency, and into arrest, implying that the joint distribution of unobserved heterogeneity will have between two (in the case of perfect correlation) and eight potential points of support. In these data, we identify six. As can be seen from panel (c) of Table 4, 41% of the sample are the type of individual that has a high susceptibility to early school leaving and a low susceptibility to both delinquency and arrest, while 30% of the sample has a high susceptibility to early school leaving and a high susceptibility to both delinquency and arrest. The next largest proportion of our sample (20%) represents the type of individual that again, has a high susceptibility to early school leaving and a high susceptibility to delinquency, although these individuals are unlikely to come into contact with the criminal justice system due to their low susceptibility to arrest. 5% of our sample represent the type of individual that has a low susceptibility to early school leaving and low susceptibilities to both delinquency and arrest. On the other hand, 1% of our sample are studious (with a low susceptibility to early school leaving) but have a high susceptibility to both delinquency and being arrested. The last type of individual (represented by 3% of the sample) is also studious and inclined toward delinquency (given a low susceptibility to early school leaving and a high susceptibility to delinquency), but is less likely to be arrested due to a low susceptibility to arrest.[16]

Panel (b) of Table 4 presents key estimates of the multivariate hazard model with the same structure as in panel (a), but for which the unobservables in the school leaving, delinquency and arrest transition rates are assumed to be independent. An LR test comparing the specifications in panels (a) and (b) confirms that the unobserved heterogeneities determining school leaving, initiation into delinquency and first arrest are not independent, with a test statistic of 542.6 exceeding the critical value of 6.0 for a χ^2_2 distribution at the 5% level of significance. Comparing the correlated model (in panel (a)) with the independent model (in panel (b)) of Table 4 shows that ignoring correlation between the unobserved heterogeneity terms has little impact on the estimate of the effect of delinquency on school leaving. Not accounting for correlation between the three transition rates results in an estimated effect of delinquency of $32\% (= 100(\exp(0.28)-1))$, compared with the correlated model, which provides an estimate of $31\% (= 100(\exp(0.27)-1))$. On the other hand, assuming independent processes results in an overestimate of the effect of arrest on school leaving of $68\% (= 100(\exp(0.52)-1))$, compared with the estimate from the

---

[16]We do not identify in the data a type of individual who has a low susceptibility to delinquency and a high susceptibility to arrest (for both high and low susceptibility to early school leaving).

correlated model of 57% ($= 100(\exp(0.45) - 1)$).

In summary, 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 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 on school leaving.

5.3 Magnitude of the effects

To demonstrate the magnitude of the effects of delinquency and arrest on school leaving, we perform simulations using the parameter estimates from our baseline model reported in panel (a) of Table 4. The simulations are for a reference male who is susceptible to early school leaving (along with 91% of our sample), is non-black and non-Hispanic, obtained a standardized (schooling-corrected) CAT-ASVAB score of 0.00, did not experience puberty prior to age 12, attended a public school, was born in 1982, did not have a teen mother at birth, does not have a responding parent (to the parent questionnaire) that is very religious, has a mother and father that are high school graduates (but have not attended college), has one younger and one older sibling, has a mother and father that are present, and resides in the suburbs, in the South.

Simulated cumulative probabilities of school leaving over the ages 14 to 25 assuming different scenarios with respect to the onset of delinquency and arrest are reported in Table 7. The first column presents the baseline case in which the reference male has not engaged in delinquency, nor been arrested. In the second and third columns the reference male is assumed to initiate delinquency at age 16, with no arrest assumed for column (2) and arrest at age 17 assumed for column (3). The fourth and fifth columns present the cumulative probability of school leaving for scenarios in which the reference male has engaged in delinquency at age 14, with no arrest assumed in column (4), and arrest at age 15 assumed in column (5).

As shown in Table 7, for the baseline case in which the reference male has not engaged in delinquency and not been arrested, the probability of leaving school at age 18 or earlier is 35%. This rises to 42% if he initiates delinquency at age 16 and is not arrested, and 55% if he has been delinquent at age 16 and arrested at age 17. Being delinquent at age 16, therefore, increases the probability of school leaving by age 18 by 20%, while subsequently being arrested at age 17 increases it by a further 37%.

Relative to delinquency, it is clear that arrest has a larger effect on an individual's probability of leaving school by age 18; however, fewer individuals have been arrested by age 17 compared with the percentage of individuals who have engaged in delinquency and not been arrested. Specifically, in our sample 17% of the sample have ever been arrested and are still in school

at age 17 (and potentially at risk of leaving school at age 18), compared with 36% who have been delinquent, not arrested, and are still in school at age 17. Roughly speaking, the impact of arrest on an individual is twice that of delinquency, but the proportion of the sample who have initiated delinquency, not been arrested and are still in school is twice the size of the proportion of the sample who have been arrested by age 17 and are still at school. This suggests that in our sample, the overall impact of delinquency in terms of school leaving is as least as large as that of arrest.

Columns (4) and (5) report on simulations in which the onset of delinquency occurs at age 14, and in the case of column (5), arrest occurs at age 15. Comparing columns (2) and (4) these simulations show that, assuming no arrest, the impact on school leaving of initiating delinquency at age 14 is very similar to the impact of initiating delinquency at age 16. However, being arrested at age 15 results in a substantial increase in the rate of school leaving at ages 16 and 17 relative to the rate of school leaving that occurs at these ages if first arrest occurs at age 17. For example by age 17, 19% of those who were arrested at age 15 have left school compared with 12% if first arrest does not occur until age 17. The gap is somewhat reduced at age 18 (58% versus 55%), and has almost been eliminated by age 19. This demonstrates that the age at which arrest first occurs has a substantial impact on early school leaving, with the earlier that arrest first takes place, the fewer grades completed. In contrast, the impact of delinquency on school leaving is smaller and its impact on school leaving differs less by age of initiation.

To gain an additional sense of the magnitude of the effects of delinquency and arrest on school leaving, we compare them to the effects of other determinants of school leaving. Specifically, looking at the results of our baseline model reported in panel (a) of Table 4, we compare the coefficient on delinquency with the coefficients on other explanators. We note that the effect of delinquency on school leaving is identical to the effect of reducing an individual's adjusted CAT-ASVAB (ability) score by one standard deviation. The coefficient on delinquency (0.27) is also similar in magnitude to that on having a mother (-0.24) or father (-0.25) present in the household. On the other hand, arrest has an effect on school leaving that is similar in magnitude (although not quite as large) as the effect on school leaving of having a father with at least a college degree (compared to having a father who has not completed high school).

6   Sensitivity analyses

In this section, we examine the sensitivity of our results, firstly, to accounting for potential reverse causality in the relationships between delinquency and school leaving, and arrest and school leaving. Next, we test the robustness of our results to allowing for an effect of prior

delinquency on the transition into arrest. We then explore whether potential differential reporting of delinquency and arrest by race influences our results. Related to this issue, we examine whether and to what extent our results are sensitive to limiting the analysis to the representative subsample of the NLSY97 (by excluding the oversample of Hispanics and blacks).

6.1 Reverse causality

As previously discussed, both common unobserved confounding factors and reverse causality are issues that must be addressed in order to identify plausibly causal effects of delinquency and arrest on school leaving. So far, we have focused on addressing the issue of common confounders. 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 (prior to age 18), typical (at age 18 or 19), and late (after age 19). Specifically, we augment our baseline specification reported in panel (a) of Table 4 by including in the starting rates for delinquency and arrest an indicator function, $I(t_s \leq 17 < t)$, equal to one if the individual leaves school before age 17 and this occurs before the current period, along with similar indicators for those who leave school at ages 18 or 19, $I(18 \leq t_s \leq 19 < t)$, and those who leave school at age 20 or older, $I(20 \leq t_s < t)$. The hazard rates for the onset of delinquency and first arrest (equations (10) and (4), respectively), therefore, become

$$
\begin{aligned}
\theta_d(t|x,t_s,u) &= \lambda_d(t)\exp(x'\beta_d + \delta_s^e I(t_s \leq 17 < t) + \delta_s^a I(18 \leq t_s \leq 19 < t) \\
&\quad + \delta_s^l I(20 \leq t_s < t) + u)
\end{aligned} \tag{14}
$$

$$
\begin{aligned}
\theta_a(t|x,t_s,\varepsilon) &= \lambda_a(t)\exp(x'\beta_a + \gamma_s^e I(t_s \leq 17 < t) + \gamma_s^a I(18 \leq t_s \leq 19 < t) \\
&\quad + \gamma_s^l I(20 \leq t_s < t) + \varepsilon)
\end{aligned} \tag{15}
$$

where t_s is the age at which an individual leaves school.[17] The parameters δ_s^e, δ_s^a and δ_s^l measure the effects of early, typical (average) and late school leaving on initiation into delinquent behavior, respectively, with the comparison category being those that do not leave school prior to initiating delinquency. Similarly, γ_s^e, γ_s^a and γ_s^l provide the effects of early, typical (average) and late school leaving on first arrest, respectively, with the comparison category being those that do not leave school prior to their first arrest.

---

[17]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.

In panel (b) of 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. Panel (a) repeats the baseline results reported in panel (a) of Table 4 for ease of reference. An LR test comparing the specifications in panels (a) and (b) confirms the existence of reverse causality, with a test statistic of 56.4 exceeding the critical value of 12.6 for a χ_6^2 distribution at the 5% level of significance.

The results reported in panel (b) 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. And while we find that the coefficient on late school leaving (after age 19) 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% percent level. In terms of magnitude, the coefficient estimate implies that early leaving school leaving (at or before age 17) increases the transition into arrest by 164% ($= 100(\exp(0.97) - 1)$, while leaving school at a typical age (18 or 19) increases it by 43% ($= 100(\exp(0.36) - 1)$. The transition rate into arrest is, therefore, higher for those who leave school earlier.

A comparison of the point estimates in panels (a) and (b) reveals that accounting for reverse causality has no impact on the estimated effect of delinquency on school leaving, and little effect on the estimated impact of arrest on school leaving (increasing it from $57\% = 100(\exp(0.45) - 1)$ to $62\% = 100(\exp(0.48) - 1)$), suggesting that the baseline results are robust 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, 2009; Buonanno and Leonida, 2009; Anderson, 2014). In contrast, we find that early school leaving has no effect on initiation into delinquency. Why is this the case? A possible explanation may be that while leaving school early does not lead to initiation into delinquency, it may lead to an increase in the intensity or severity of offending amongst those who have previously initiated, and it is this increase in intensity and severity that leads to arrest. We attempt to investigate whether school leaving increases the intensity and severity of delinquent behavior using (panel) data on the number of delinquent acts committed and on the income earned from delinquency in the year prior to and the year after leaving school.[18] We find that,

---

[18]We use information from the NLSY97 on the intensity of delinquency and on income from delinquency. The intensity measure is constructed using individual information on the number of times that the respondent engaged in each type of delinquent activity (destroyed property, stolen more than $50 worth, other property crimes, attack, and sold drugs) in the last 12 months in wave 1, and in subsequent waves, since the date of last interview. If the respondent stole something worth $50 or more, committed other property crimes, or sold or helped to sell drugs,

of those who engage in delinquency in the year before they leave school, 44% report committing 10 or more delinquent acts, while 79% of those who report any income from delinquency in the year before leaving school report receiving $100 or more. Within this group, 47% of those who engage in delinquency in the year after leaving school commit 10 or more offenses, and 86% of those reporting an income from crime the year after leaving school receive $100 or more. This evidence is consistent with early school leaving increasing the intensity and severity of delinquent behavior amongst delinquents. It seems reasonable that this, in turn, leads to their arrest.

6.2 Effect of delinquency on arrest

In panel (c) of Table 5, we investigate the effect of delinquency on an individual's transition into arrest by including in the arrest equation an indicator for prior delinquency. An LR test comparing the specifications in panels (a) and (c) confirms that prior delinquency impacts the transition into arrest, with a test statistic of 9.0 exceeding the critical value of 3.8 for a χ^2_1 distribution at the 5% level of significance. Prior delinquency, is found to increase the transition rate into arrest by 28% ($= 100(\exp(0.25)-1)$). Importantly, accounting for the effect of initiating delinquency on arrest has no effect on the estimate of the effect of arrest on school leaving (which remains at 57% ($= 100(\exp(0.45)-1)$ in both panels (a) and (c)), and little effect on the estimate of the effect of delinquency on school leaving, which goes from 31% ($= 100(\exp(0.27)-1)$ in panel (a) to 30% ($= 100(\exp(0.26)-1)$ in panel (d). From this we conclude that our main findings are robust to accounting for the effect of delinquency on the transition into arrest.

In panel (d) of Table 5, we investigate whether an individual's transition into arrest differs according to the age at which he first engages in delinquency. To do so, we expand the baseline model in panel (a) by including in the arrest equation an indicator for whether the individual initiates delinquency at or before age 17, and an indicator for initiation at age 18 or older (given that at 18 individuals have reached the age of majority). However, an LR test comparing the specifications in panels (c) and (d) does not provide evidence in favor of differential effects of delinquency according to age of initiation, with a test statistic of 1.2 being lower than the critical value of 3.8 for a χ^2_1 distribution at the 5% level of significance.[19]

he was asked (in each round) about the total cash received (or the total cash he would have received) from selling these items in the last year, and this is used to construct the income measure. Due to missing information we are able to establish the intensity and severity of delinquent activity for a limited number of individuals. Specifically, we are able to ascertain the number of delinquent acts committed by 223 individuals in the year prior to and the year after school leaving, and we are able to establish the income arising from delinquency for 80 individuals in the year prior to and the year after leaving school.

[19]Similarly, in an unreported specification in which we expand the baseline model in panel (a) by including in the arrest equation an indicator for whether the individual initiates delinquency at or before age 15, an indicator for whether delinquency was initiated at age 16 or 17, and an indicator for initiation at age 18 or older, an LR test once again fails to provide evidence in favor of differential effects of delinquency according to age of initiation on the transition into arrest.

6.3 Differential reporting by race

As mentioned in Section 3.4, previous research suggests there may exist differential tendencies to report delinquent behavior by race (Hindelang et al., 1981; Lochner and Moretti, 2004). In order to investigate the potential impact of differential reporting by race on our findings, we expand our baseline model to include interaction terms between delinquency and race (black, Hispanic, with non-Hispanic and non-black as the comparison category) and between arrest and race. Evidence of statistically significant interaction terms may suggest differential reporting. The results are presented in panel (e) of Table 5. An LR test comparing the specifications in panels (a) and (e) confirms the interaction terms are jointly significant, with a test statistic of 16.0 exceeding the critical value of 9.5 for a χ_4^2 distribution at the 5% level of significance.

Turning to the coefficient estimates in panel (e), we note that the interaction terms between the race indicators and arrest are not individually statistically significant. The common effect of arrest across race suggests an absence of differential reporting of arrest. In the case of delinquency, however, we do find negative and statistically significant interaction terms with the indicators for race. Interpreting the magnitude of the effect of delinquency by race, we find that delinquency has the largest effect on school leaving for those who are non-black/non-Hispanic $(49\% = 100(\exp(0.40) - 1))$, with smaller effects for blacks $(13\% = 100(\exp(0.40 - 0.28) - 1))$ and Hispanics $(6\% = 100(\exp(0.40 - 0.34) - 1))$.

Comparing the results in panels (a) and (e) of Table 5, it is clear that the coefficient on delinquency for the baseline model is most similar to (and smaller in magnitude than) that for non-black/non-Hispanic people. Specifically, the coefficient on delinquency is 0.27 in the baseline model in panel (a) compared with 0.40 for non-black/non-Hispanics in panel (e). 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.

We further explore this issue by removing from the estimation sample the "oversample" of Hispanics and blacks and limiting our analysis to the representative cross-sectional sample (N=3,393) .[20] The results from doing so are reported in panel (f) of Table 5. As expected, they show that using only the cross-sectional sample leads to small increases in the magnitudes of the estimated effects of delinquency and arrest on school leaving (which remain statistically significant at the 1% level). Specifically, the effect of delinquency on school leaving increases from $31\% (= 100(\exp(0.27) - 1)$ in panel (a) to $38\% (= 100(\exp(0.32) - 1)$ in panel (f). Similarly, in going from panel (a) to (f), the effect of arrest on school leaving increases from 57% (=

---

[20]The NLSY97 includes two subsamples; the first is a cross-sectional sample representative of people living in the United States, while the second oversamples Hispanic and black people.

$100(\exp(0.45) - 1)$ to 60% $(= 100(\exp(0.47) - 1))$. We conclude that our results are robust to using only the cross-sectional sample of the NLSY97.

7   Mechanisms

7.1   Income versus 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). 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. In order 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 generating delinquency and into non-income generating delinquency, and allow each to affect the school leaving rate.[21] Combined with the hazards for arrest and school leaving, this results in a four equation system of hazards in which the four processes have a correlated error structure.

The results from maximum likelihood estimation of the four equation system of hazards are presented in Table 6. The first column provides estimates for the income delinquency equation, the second column for the non-income delinquency equation, the third column for the arrest equation, and the fourth column for the school leaving equation. Panel (a) of Table 6 reports the parameter estimates assuming that the unobserved heterogeneity terms in the four equations are correlated, panel (b) reports key estimates assuming the unobserved heterogeneities are independent, and panel (d) reports the joint distribution of unobserved heterogeneity implied by the estimates in panel (a). The parameters of interest are the coefficients on arrest, income and non-income delinquency in the equation for school leaving.

As can been seen from panel (a) of Table 6, after accounting for correlated unobserved heterogeneity, 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 coefficient on non-income generating delinquency is both smaller in magnitude and weaker in terms of significance. The estimates imply that the school leaving rate is increased by 49% ($100(\exp(0.40)$-1)) following arrest, and by 41% ($100(\exp(0.34)$-1)) following the transition into income generating delinquency.

---

[21]Income generating delinquent acts include stealing more than \$50 worth, other property crimes, and selling drugs. Non-income generating delinquent acts include destroying property and attacking others.

Non-income generating delinquency increases the school leaving rate by just 9% (100(exp(0.09)-1)). These results are consistent with the prediction from the capital accumulation model of crime that the effect of income generating delinquency on school leaving is larger than that of non-income producing delinquency. In addition, the results find that the magnitudes of the effects of arrest and income generating delinquency on school leaving are similar.

In terms of unobserved heterogeneity, panel (a) shows that we identify two types in each of the transition rates for income and non-income generating delinquency, as well as two types in each of the transition rates for arrest and school leaving. This suggests that there are at least 2 (in the case of perfect correlation) and at most 16 types of individuals in the joint distribution of unobserved heterogeneity. We identify 11 types in these data. The joint distribution of unobserved heterogeneity is presented in panel (d) of Table 6 which shows, conditional on observed characteristics, the percentage of the sample that represents each type of individual. As before, most of our sample (43%) are the type with a high susceptibility to early school leaving and a low susceptibility to (both types of) delinquency and arrest. The type represented by the next highest proportion of the sample (22%) also has a high susceptibility to early school leaving; however, this type has a high susceptibility to both types of delinquency and arrest. The next two most common types in our data (representing 11% and 10% of the sample, respectively) have a high susceptibility to early school leaving and a high susceptibility to non-income generating delinquency either alone or in combination with income generating delinquency; however, these types are less likely to be punished due to their low susceptibility to arrest.[22]

Panel (b) of Table 6 presents estimates for the specification of the four equation system of hazards in which the unobservables in each equation are assumed to be independent across equations. Using an LR test to compare the specifications in panels (a) and (b) confirms that the unobserved heterogeneities determining school leaving, arrest, income and non-income delinquency are not independent, with a test statistic of 1668.4 exceeding the critical value of 14.1 for a χ^2_7 distribution at the 5% level of significance. Comparing the correlated model (in panel (a)) with the independent model (in panel (b)) of Table 6 shows that ignoring the correlation between the unobserved heterogeneity terms results in an overestimate of the effect of arrest on school leaving of 60% (= 100(exp(0.47) − 1)), compared with the estimate from the correlated model of 49% (100(exp(0.40)-1)).

In summary, we find that the causal effect of initiation into delinquency on school leaving is primarily driven by initiation into income generating delinquency. Given that income-generating

---

[22]We do not identify in our sample a type that combines a high susceptibility for arrest with a low susceptibility for both income and non-income generating delinquency. Although we identify the type of individual that has a high susceptibility to non-income generating delinquency and arrest, along with a low susceptibility to income generating delinquency, the percentage of the sample represented by this type is so small that it is rounded to zero.

delinquency can be viewed as a (risky) occupation offering favorable monetary returns relative to legitimate work, these results are consistent 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 affects school leaving. 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.

7.2 Age at initiating delinquency and first arrest

The four equation system of hazards presented in panel (a) of Table 6 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 order to investigate this issue, we expand the model in panel (a) of Table 6 by allowing the effects of (income and non-income generating) delinquency and arrest to vary according to whether they first occur at an age younger than 15 (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). The key results are presented in panel (c) of Table 6. An LR test comparing the specifications in panels (a) and (c) confirms the existence of differential age of onset effects for delinquency and arrest on school leaving, with a test statistic of 21.0 exceeding the critical value of 12.6 for a χ_6^2 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, while the indicator for first arrest occurring at 18 or older is not. In terms of the magnitude of effects, first arrest by age 15 increases the school leaving rate relative to an otherwise similar individual who is not arrested by 72% ($= 100(\exp(0.54) - 1)$), whereas first arrest at ages 16 or 17 increases it by 38% ($= 100(\exp(0.32) - 1)$).[23] The school leaving rate is, therefore, higher for those who are first arrested earlier.

Similarly, with respect to income generating delinquency we find that while initiation at

---

[23]Although not statistically different from zero at conventional levels, the magnitude of the point estimate implies that arrest at age 18 or older increases the school leaving rate by 15% ($= 100(\exp(0.14) - 1)$).

all ages has a statistically significant effect on school leaving, the effect is largest for initiation that occurs early (by age 15). Compared to an otherwise similar individual who does not engage in income generating delinquency, initiation into income delinquency by age 15 increases the school leaving rate by 45% ($= 100(\exp(0.37) - 1)$), initiation at ages 16 or 17 by 34% ($= 100(\exp(0.29) - 1)$), while initiation at ages 18 or older increases the school leaving rate by 38% ($= 100(\exp(0.32) - 1)$). [24] Overall, these results and those for first arrest are consistent with predictions from a capital accumulation model of crime, suggesting the accumulation of criminal capital as a potential mechanism via which delinquency 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. Specifically, initiation into non-income delinquency at age 18 or older increases the school leaving rate by 31% ($= 100(\exp(0.27) - 1)$) compared to an otherwise similar individual who does not engage in non-income generating delinquency. 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.[25] This may occur, for example, if violent behavior 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.

8 Discussion

This paper provides empirical evidence on the role of delinquency and arrest in determining early school leaving. To do so, a multivariate mixed proportional hazard framework is used to model the transitions into delinquency, arrest, and school leaving as a system. Using data on males from the National Longitudinal Survey of Youth 1997, and accounting for common unobservable confounders and reverse causality, we find that both first arrest and initiation into delinquency lead to early school leaving. In terms of the magnitude of estimated effects, being arrested for the first time is found to have roughly twice the effect of initiation into delinquency on the likelihood that an individual leaves school. However, amongst males in our sample who are enrolled in school at age 17, twice as many have been delinquent but not arrested compared to the number arrested. Therefore, the overall impact of delinquency on school leaving is at least as large as that of arrest in our sample.

---

[24] 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), with a $\chi_1^2 = 0.2$.

[25] The 1984 dropout rate (being the proportion of a given age cohort that is not enrolled in school and has not completed high school) for persons aged 16 and 17 was 6.8%, while it was 15.2% for those aged 18 and 19 (Rumberger, 1987).

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 consistent 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 salient or vulnerable ages with regard to the school leaving decision. In terms of reverse causality, we find that leaving school early does increase the likelihood of first arrest, but only if it occurs prior to completing high school (by age 19) with larger effects when first arrest occurs earlier. Finally, we find no evidence that initiation into delinquency leads to early school leaving.

Our findings provide useful insights for policy development. Specifically, we show that there are a large group of delinquents who avoid arrest, but whose reduced level of educational attainment is as important as that of the group that have been arrested. As a result, to focus interventions solely on those who come to the attention of the criminal justice system would miss a large part of the vulnerable population. 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 interventions targeting this group may be most effectively implemented within schools. 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 neighborhoods. 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.[26] Further examples of such programs include "Functional Family Therapy", and the "PATHS (Promoting Alternative Thinking Strategies) Curriculum".

The second policy relevant finding from this 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 dis-

---

[26]The intervention included regular interactions with a pro-social adult, after-school programming and in-school programming aimed at recognising and reducing unhelpful, automatic thinking and judgment errors through cognitive behavioral therapy (CBT).

advantaged individuals who may not otherwise have access to reliable information, may reduce early school leaving (Jensen, 2010).[27] 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 labor 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 toward schooling increase individuals' expected returns to education (Heckman, 2000; Johnson, 1997).[28]

Finally, our research highlights the complex and bi-directional nature of the relationship between delinquency, arrest and education. While we narrow our focus to consider the impacts of initiation into delinquency and first arrest on school leaving, intensity and persistence of offending and arrest are also important dimensions to consider in understanding the delinquency, arrest, schooling relationship. Pursuing these issues is an important area for future research.

---

[27]This information may have a larger impact if returns to education in the long and short run are contrasted with returns to crime. However, doing so may raise ethical questions and so not be possible in practice.

[28]The Philadelphia Futures' Sponsor-A-Scholar (SAS) program seeks to help public high school students continue their education beyond high school to college by providing mentoring throughout high school and for a year afterward, academic support, assistance in applying to college and for financial aid, and financial assistance throughout college.

References

Abbring, J. H. and van den Berg, G. J. (2003). The nonparametric identification of treatment effects in duration models. *Econometrica*, 71(5):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*, 130(2):759 – 803.

Anderson, D. A. (1999). The aggregate burden of crime. *Journal of Law and Economics*, 42(2):611 – 642.

Anderson, D. M. (2014). In school and out of trouble? The minimum dropout age and juvenile crime. *Review of Economics and Statistics*, 96(2):318 – 331.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):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*, 95(1-2):41 – 53.

Bulow, W. (2014). The harms beyond imprisonment: Do we have special moral obligations towards the families and children of prisoners? *Ethical Theory and Moral Practice*, 17(4):775 – 789.

Buonanno, P. and Leonida, L. (2009). Non-market effects of education on crime: Evidence from Italian regions. *Economics of Education Review*, 28(1):11 – 17.

Carneiro, P., Heckman, J. J., and Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, 101(6):2754 – 2781.

Carneiro, P., Meghir, C., and Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, 11:123 – 160.

Carson, E. A. (2014). *Education and correctional populations*. US Department of Justice, Bureau of Justice Statistics, Washington, DC.

Christian, J., Mellow, J., and Thomas, S. (2006). Social and economic implications of family connections to prisoners. *Journal of Criminal Justice*, 34(4):443 – 452.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy*, 81(3):521.

Elliott, D. S. and Voss, H. L. (1974). *Delinquency and dropout.* DC Heath, Lexington, MA.

Farrington, D. P. (1973). Self-reports of deviant behavior: Predictive and stable? *Journal of Criminal Law and Criminology*, 64(1):99 – 110.

Federal Bureau of Investigation (2012). *Crime in the United States 2012.* Uniform Crime Reports. Washington, DC.

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.*

Grogger, J. (1998). Market wages and youth crime. *Journal of Labor Economics*, 16(4):756 – 791.

Hansen, K. T., Heckman, J. J., and Mullen, K. J. (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics*, 121(1/2):39 – 98.

Harlow, C. W. (2003). *Education and correctional populations.* US Department of Justice, Bureau of Justice Statistics, Washington, DC.

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, 54(1):3 – 56.

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*, 24(3):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(2):613 – 630.

Jacob, B. A. and Lefgren, L. (2003). Are idle hands the devil's workshop? Incapacitation, concentration, and juvenile crime. *American Economic Review*, 93(5):1560 – 1577.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, 125(2):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. *Paper AAI9727243. UPenn dissertation available from ProQuest.*

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.

Kirk, D. S. and Sampson, R. J. (2013). Juvenile arrest and collateral educational damage in the transition to adulthood. *Sociology of Education*, 86(1):36 – 62.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83 – 119.

Kupchik, A. (2010). *Homeroom security: School discipline in an age of fear.* New York University Press, New York.

Lehrer, S. F., Tremblay, R. E., Vitaro, F., and Schaal, B. (2004). Raging hormones in puberty: Do they influence adolescent risky behavior? *Working Paper, Queen's University.*

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

Lochner, L. (2004). Education, work, and crime: A human capital approach. *International Economic Review*, 45(3):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*, 94(1):155 – 189.

Luallen, J. (2006). School's out... forever: A study of juvenile crime, at-risk youths and teacher strikes. *Journal of Urban Economics*, 59(1):75 – 103.

Machin, S., Marie, O., and Vujić, S. (2011). The crime reducing effect of education. *Economic Journal*, 121(552):463 – 484.

Machin, S., Marie, O., and Vujić, S. (2012). Youth crime and education expansion. *German Economic Review*, 13(4):366 – 384.

Machin, S. and Meghir, C. (2004). Crime and economic incentives. *Journal of Human Resources*, 39(4):958 – 979.

Merlo, A. and Wolpin, K. (2009). The transition from school to jail: Youth crime and high school completion among black males. *Working Paper, Penn Institute for Economic Research.*

Mocan, N. and Rees, D. I. (2005). Economic conditions, deterrence and juvenile crime: evidence from micro data. *American Law and Economics Review*, 7(2):319 – 349.

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

Moffitt, T. E. (2006). Life-course-persistent versus adolescence-limited antisocial behavior. In Cicchetti, D. and Cohen, D. J., editors, *Developmental psychopathology: Risk, disorder, and adaptation*, volume 3, pages 570 – 598. John Wiley and Sons Inc, Hoboken, NJ, 2nd edition.

Mustard, D. B. (2010). How do labor markets affect crime? New evidence on an old puzzle. *No. 4856. Discussion paper series, Institute for the Study of Labor (IZA).*

Neumark, D. and Joyce, M. (2001). Evaluating school-to-work programs using the new NLSY. *Journal of Human Resources*, 36(4):666 – 702.

Nottelmann, E. D., Susman, E. J., Inoff-Germain, G., Cutler, G. B., Loriaux, D. L., and Chrousos, G. P. (1987). Developmental processes in early adolescence: Relationships between adolescent adjustment problems and chronological age, pubertal stage, and puberty-related serum hormone levels. *Journal of Pediatrics*, 110:473 – 480.

Office of Juvenile Justice and Delinquency Prevention (2013). *Census of Juveniles in Residential Placement 1997-2011.* Washington, DC.

Oreopoulos, P. and Salvanes, K. G. (2009). How large are returns to schooling? Hint: money isn't everything. *No. 15339. National Bureau of Economic Research.*

Rumberger, R. W. (1987). High school dropouts: A review of issues and evidence. *Review of Educational Research*, 57(2):101 – 121.

Sampson, R. J. and Laub, J. H. (2005). A life-course view of the development of crime. *Annals of the American Academy of Political and Social Science*, pages 12 – 45.

Schwartz-Soicher, O., Geller, A., and Garfinkel, I. (2011). The effect of paternal incarceration on material hardship. *Social Service Review*, 85(3):447 – 473.

Segal, C. (2013). Misbehavior, education, and labor market outcomes. *Journal of the European Economic Association*, 11(4):743–779.

Webbink, H. 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*, 29(1):115 – 144.

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*, 37(3):479 – 509.

Appendix A: Data Construction

A.1 Outcomes

A.1.1 Delinquency

The age at which an individual initiates delinquent behavior is constructed by combining information collected in wave one of the NLSY97 on the age at which respondents first committed a set of listed delinquencies, and information collected in subsequent waves on whether they had committed the same set of listed offences since the last interview. The listed delinquencies are: (1) purposely damaged or destroyed property that did not belong to the respondent; (2) stolen something from a store, person or house, or something that did not belong to the respondent worth $50 or more including stealing a car; (3) committed other property crimes such as fencing, receiving, possessing or selling stolen property, or cheated someone by selling them something that was worthless or worth much less than what the respondent said it was; (4) attacked someone with the idea of seriously hurting them or have a situation end up in a serious fight or assault of some kind; and (5) sold or helped sell marijuana (pot, grass), hashish (hash) or other hard drugs such as heroin, cocaine or LSD. This detailed information allows us to distinguish between the age at which an individual commits his first income generating delinquent act and the age at which he commits his first non-income generating act. Income generating delinquent acts are stealing something worth $50 or more, other property crimes, and selling drugs. Non-income generating acts are attacking someone, and destroying property. The NLSY97 asks the delinquency questions of all the respondents up until round 8 in 2004 (when individuals are aged 19–25). From round 8 onwards, the delinquency questions are only asked of respondents who had ever reported being arrested and a control group of approximately 10% of the respondents for comparison. This feature of the data is naturally handled with a duration framework in which incomplete observations are treated as right censored.

A.1.2 Arrest

The age at which an individual is first arrested is constructed using information from all thirteen waves of the NLSY97. Specifically, the first wave of the NLSY97 asks each respondent whether he had ever been arrested by the police or taken into custody for an illegal or delinquent offence (not including arrests for minor traffic violations), and if so, his age at the time of his first arrest. The survey then follows up each year by asking whether the respondent has been arrested by the police or taken into custody for an illegal or delinquent offence (not including arrests for minor traffic violations) since the date of last interview. We combine this with information compiled by

the NLSY97 across all waves providing the earliest arrest date reported by the respondent.[29,30]

A.1.3 School Leaving

In order to create the age at which a respondent initiates delinquency and is first arrested, information is available on whether the respondent has been delinquent or arrested between two interview dates with the survey asking, for example, whether the respondent had been arrested since the date of last interview. We, therefore, assign the age of first arrest or delinquency as the age at the date of interview when the respondent first answers affirmatively. More accurate information is available to construct the age of an individual's first school leaving, such as an individual's monthly enrolment status. However, as the timing of our three outcomes relative to one another is critical to the accuracy of our econometric approach, we ensure consistency by using the same approach to assign timing to school leaving as we did for arrest and delinquency. Specifically, we choose to use information provided by the survey that allows us to ascertain whether the individual became unenrolled between two interview dates, assigning the age of school leaving as the age at the date of interview when the respondent first becomes unenrolled.

The age at which a respondent first leaves school (defined as primary, secondary or tertiary education) is, therefore, constructed using information collected on respondents' most recent grade of school completed (or the number of years spent at college) and their educational institution enrolment status at the time of each interview.[31,32] If the individual leaves school prior to the first wave (when individuals are aged 12–18), we use the age at which the respondent reports leaving high school.

---

[29]The age of criminal responsibility in the U.S. is between 6 and 12 depending on the state. However, less than 0.1% of the arrests made in 2012 were of individuals aged less than 10 (Federal Bureau of Investigation, 2012). We, therefore, treat those who report earlier arrest as having been arrested for the first time at age 10. There are 32 individuals who report having been arrested prior to age 10 that are recoded as a result.

[30]Included in our sample are 266 individuals who report being arrested for the first time without reporting having been delinquent. There are also 222 individuals who report initiating delinquency after their first arrest. There are at least two explanations for this. The first is that the respondent may have been falsely arrested. The second is that the individual may have been arrested for an activity that is not captured by our measure of delinquency. For example, an individual may have been arrested for protesting, an activity that the NLSY97 did not include when asking respondents about having committed a set of listed delinquencies. Nevertheless, our findings are robust to dropping these individuals from the sample.

[31]The enrolment variable considers those working toward a GED as being enrolled regardless of where that course of study took place.

[32]Given that we define our measure to be the *first* time a respondent leaves school, a potential concern is that the effect of delinquency and arrest on school leaving might differ if individuals subsequently return to school. However, we estimate a probit model of an indicator for returning to school as a function of an indicator for delinquency prior to first school leaving, an indicator for arrest prior to first school leaving, and our set of regressors. We find that the effect of prior delinquency on returning to school is negative and statistically insignificant, while the effect of prior arrest on returning to school is also negative but significant. Similar results are obtained using a linear probability model.

A.2   Explanatory Variables

Information is taken from the first round of the NLSY97 (when individuals were aged 12–18), unless otherwise stated.

A.2.1   Individual characteristics

- Race: indicator variables – Hispanic; black; non-black and non-Hispanic (as the base category).

- Adjusted CAT-ASVAB (ability) score: standardized schooling-corrected CAT-ASVAB score with a mean of zero and a standard deviation of one[33] – in 1999, 79.3 percent of the NLSY97 sample respondents were paid \$75 each to complete the computer-adaptive form of the Armed Services Vocational Aptitude Battery (CAT-ASVAB), for which each person achieved a score between zero and 100. The CAT-ASVAB score is a summary of four tests taken: Mathematical Knowledge, Arithmetic Reasoning, Word Knowledge and Paragraph Comprehension. The score is likely to be affected by schooling, so we correct it for educational attainment in the year of the test.[34]

- Puberty before age 12: indicator variable – each year, the NLSY97 asks each respondent whether he has experienced any biological changes associated with puberty (such as the development of facial or pubic hair, or the cracking or lowering of his voice) and if so, how old he was when these changes began.[35]

- Private/parochial school attendance: indicator variable – based on whether the respondent's current or most recent elementary, middle or high school attended was public, private, or parochial.

- Year of birth: indicator variables – 1980, 1981, 1982, 1983, 1984 (as the base category).

---

[33]It is common in the literature to use test scores determining an individual's eligibility for service in the armed forces as a proxy for cognitive skills (Heckman et al., 2006).

[34]To correct the CAT-ASVAB score for schooling, we use a method based on Hansen et al. (2004) to estimate the effect of schooling on the test score. This procedure exploits the randomness in the amount of schooling undertaken at the test date, since all the respondents took the CAT-ASVAB in the same year. Our schooling-corrected test score is, therefore, the residual of a regression of the CAT-ASVAB score on years of education completed in the year of the test, holding final completed schooling constant (Carneiro et al., 2011, 2013).

[35]This variable allows us to determine the effects of entering puberty at an age younger than usual. Biological changes during puberty mean that adolescents begin producing hormones such as testosterone and cortisol. For example, during puberty, males experience an 18 fold increase in testosterone levels on average (Nottelmann et al., 1987). These hormones have strong effects on many tissues of the body, such as the brain (Lehrer et al., 2004). Lehrer et al. (2004) find that there are significant effects of hormone levels on many risky and criminal activities during puberty. We, therefore, anticipate that the younger an individual when he enters puberty, the more likely he would be to engage in risky activities such as delinquent behavior at a young age.

A.2.2 Parental characteristics

- Mother's age at respondent's birth ≤ 19: indicator variable.

- Responding parent very religious: indicator variable equal to one if the responding parent has a religiosity score greater than four, based on a score between zero and six, with zero being 'not religious' and six being 'very religious'; obtained from the parent questionnaire administered in round 1.

- Father's (Mother's) education: indicator variables – high school graduate, more than high school but less than college, college graduate or more, less than a high school graduate (as the base category); separate sets of dummy variables are created for the respondent's mother and father; based on information from the parent questionnaire administered in round 1 and the childhood retrospective section introduced in wave 6 to obtain information from respondents without a completed parent questionnaire.

A.2.3 Household characteristics

- Number of younger (older) siblings: based on ages of siblings (relative to the respondent) recorded in the household and non-household rosters of the respondent's relatives.

- Mother (Father) present in household: indicator variables – based on the NLSY97's household roster, establishing the relationship of each person in the household to the respondent.

A.2.4 Local environment characteristics

- Resides in suburbs; city; rural area (as the base category): indicator variables – based on information indicating whether the individual resides in an urban or rural area, combined with information on whether the respondent lives in a metropolitan statistical area (MSA).[36] In specific, whether the respondent:

  - lives in an MSA but not in the central city (i.e., in the suburbs)
  - lives in an MSA in the central city (i.e., in the central business district)
  - lives in an MSA (unknown whether in the central city)[37]
  - does not live in an MSA

---

[36]A metropolitan statistical area (MSA) is a geographical region with a high core population density and close economic ties throughout the area.

[37]We allocate individuals for whom it is unknown whether they live in the central city to the suburbs, as we assume the confusion stems from living along the city's boundary.

- Resides in North; South; West; Northeast (as the base category): indicator variables.

For some of the control variables (ability, parent education, teen mother at birth and parent very religious), we have incomplete data. For example, the Armed Services Vocational Aptitude Battery (CAT-ASVAB) was administered to only 79% of the NLSY97 sample. In order to preserve the sample size, we assign the average to missing values. We therefore assign the average CAT-ASVAB score to individuals if their scores are missing (Kling et al., 2007; Heller et al., 2013).

Table 1: Descriptive Statistics

Outcomes	Mean	Minimum	Maximum
School leaving			
Left school (during observation period)	0.94	0.00	1.00
Age first left (conditional on leaving)	19.52	14.00	29.00
Delinquency			
Engaged in delinquency	0.68	0.00	1.00
Age of initiation (conditional on initiating)	12.87	7.00	29.00
Income delinquency			
Engaged in income delinquency	0.43	0.00	1.00
Age of initiation (conditional on initiating)	14.83	7.00	29.00
Non-income delinquency			
Engaged in non-income delinquency	0.61	0.00	1.00
Age of initiation (conditional on initiating)	12.92	7.00	29.00
Arrest			
Arrested	0.44	0.00	1.00
Age of first arrest (conditional on having been arrested)	17.40	10.00	29.00

Explanators			
Individual characteristics			
Race is black	0.25	0.00	1.00
Race is Hispanic	0.21	0.00	1.00
Race is non-black and non-Hispanic	0.54	0.00	1.00
Adjusted CAT-ASVAB (ability) score[a]	0.00	-1.71	2.12
Puberty before age 12	0.26	0.00	1.00
Private/parochial school attendance	0.06	0.00	1.00
Year of birth: 1980	0.19	0.00	1.00
Year of birth: 1981	0.21	0.00	1.00
Year of birth: 1982	0.21	0.00	1.00
Year of birth: 1983	0.20	0.00	1.00
Year of birth: 1984	0.20	0.00	1.00
Parental characteristics[a]			
Mother's age at respondent's birth ≤ 19	0.13	0.00	1.00
Responding parent very religious	0.36	0.00	1.00
Father: < high school graduate	0.24	0.00	1.00
Father: high school graduate	0.39	0.00	1.00
Father: >HS & <college grad	0.17	0.00	1.00
Father: $\geq$college graduate	0.20	0.00	1.00
Mother: < high school graduate	0.23	0.00	1.00
Mother: high school graduate	0.37	0.00	1.00
Mother: >HS & <college grad	0.22	0.00	1.00
Mother: $\geq$college graduate	0.18	0.00	1.00
Household characteristics			
Number of younger siblings	1.19	0.00	8.00
Number of older siblings	0.97	0.00	11.00
Father present in the household	0.69	0.00	1.00
Mother present in the household	0.92	0.00	1.00
Local environment characteristics			
Resides in suburbs	0.36	0.00	1.00
Resides in city	0.31	0.00	1.00
Resides in rural area	0.33	0.00	1.00
Resides in Northeast	0.18	0.00	1.00
Resides in North	0.23	0.00	1.00
Resides in South	0.37	0.00	1.00
Resides in West	0.22	0.00	1.00
Observations	4,488		

[a] $N < 4,488$

Table 2: First delinquency, arrest and school leaving – timing of events (percentages)

Order of events	Delinquency						Arrest	
	(a) Any		(b) Income		(c) Non-income		(d)	
Delinquency/arrest, school leaving	2811	(62.7)	1650	(36.8)	2548	(56.8)	1206	(26.8)
School leaving, delinquency/arrest	141	(3.1)	180	(4.0)	131	(2.9)	572	(12.8)
Delinquency/arrest = school leaving	94	(2.1)	114	(2.5)	81	(1.8)	182	(4.1)
No delinquency/arrest	1442	(32.1)	2544	(56.7)	1728	(38.5)	2528	(56.3)
Total	4488	(100.0)	4488	(100.0)	4488	(100.0)	4488	(100.0)

Table 3: Parameter estimates – bivariate hazard model: arrest and school leaving

	Arrest		School Leaving	
(a) Correlated unobserved heterogeneity				
Effect arrest		–	0.42	(8.7)***
Individual characteristics				
Race is black	0.08	(1.0)	-0.13	(2.4)**
Race is Hispanic	-0.18	(2.2)**	-0.05	(1.0)
Adjusted CAT-ASVAB (ability) score	-0.22	(6.1)***	-0.27	(12.2)***
Puberty before age 12	0.17	(2.7)***	0.07	(1.6)
Private/parochial school attendance	-0.24	(1.8)*	-0.36	(4.4)***
Year of birth: 1980	-0.23	(2.4)**	-0.12	(2.0)**
Year of birth: 1981	-0.05	(0.5)	-0.04	(0.7)
Year of birth: 1982	0.03	(0.3)	0.00	(0.0)
Year of birth: 1983	-0.06	(0.7)	0.00	(0.1)
Parental characteristics				
Mother's age at respondent's birth ≤ 19	0.16	(1.8)*	0.12	(2.1)**
Responding parent very religious	-0.21	(3.1)***	-0.13	(2.8)***
Father: high school graduate	-0.18	(2.3)**	-0.11	(2.0)**
Father: >HS & <college grad	-0.18	(2.0)**	-0.22	(3.6)***
Father: $\geq$college graduate	-0.58	(5.2)***	-0.60	(8.3)***
Mother: high school graduate	-0.26	(3.3)***	-0.17	(3.2)***
Mother: >HS & <college grad	-0.25	(2.8)***	-0.35	(5.6)***
Mother: $\geq$college graduate	-0.13	(1.2)	-0.60	(8.2)***
Household characteristics				
Number of younger siblings	0.03	(1.3)	0.07	(4.5)***
Number of older siblings	0.02	(0.9)	0.04	(2.1)**
Father present in the household	-0.28	(4.2)***	-0.26	(5.9)***
Mother present in the household	-0.50	(4.9)***	-0.29	(4.0)***
Local environment characteristics				
Resides in suburbs	0.14	(1.9)*	-0.09	(1.9)*
Resides in city	0.19	(2.5)**	-0.04	(0.8)
Resides in North	0.24	(2.6)***	0.09	(1.6)
Resides in South	0.11	(1.3)	0.10	(1.8)*
Resides in West	0.14	(1.5)	0.08	(1.4)
Unobserved heterogeneity				
Constant: ε_1, v_1	-3.34	(11.7)***	-5.39	(16.1)***
ε_{22}, v_{22}	-1.83	(6.1)***	-1.76	(15.1)***
α_1		1.61	(4.3)***	
α_2		-2.87	(1.1)	
α_3		1.35	(5.3)***	
$-$Loglikelihood			16782.1	
(b) Independent unobserved heterogeneity				
Effect arrest		–	0.60	(14.3)***
$-$Loglikelihood	8126.7		8667.0	
LR test statistic		23.2**		

(c) Joint distribution of unobserved heterogeneity

	Arrest			
	High	Low	Total	Independent
School leaving				
High	51	39	90	91
Low	0	10	10	9
Total	51	49	100	
Independent	26	74		100

Absolute t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The structure of age dependence for school leaving: age dummies 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+. For arrest: age dummies 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+.

Table 4: Parameter estimates – multivariate hazard model: delinquency, arrest and school leaving

	Delinquency		Arrest		School Leaving	
(a) Correlated unobserved heterogeneity						
Effect delinquency	–		–		0.27	(5.5)***
Effect arrest	–		–		0.45	(9.0)***
Individual characteristics						
Race is black	-0.14	(1.8)*	0.06	(0.7)	-0.13	(2.4)**
Race is Hispanic	-0.44	(5.5)***	-0.18	(1.9)*	-0.03	(0.5)
Adjusted CAT-ASVAB (ability) score	-0.12	(3.9)***	-0.26	(6.8)***	-0.27	(11.9)***
Puberty before age 12	0.29	(4.9)***	0.21	(2.8)***	0.04	(1.0)
Private/parochial school attendance	-0.22	(1.9)*	-0.32	(2.1)**	-0.33	(4.1)***
Year of birth: 1980	-0.44	(4.9)***	-0.40	(3.6)***	-0.11	(1.8)*
Year of birth: 1981	-0.27	(3.1)***	-0.09	(0.9)	-0.05	(0.8)
Year of birth: 1982	-0.09	(1.1)	-0.04	(0.4)	-0.02	(0.3)
Year of birth: 1983	-0.18	(2.2)**	-0.13	(1.3)	0.01	(0.2)
Parental characteristics						
Mother's age at respondent's birth ≤ 19	-0.04	(0.5)	0.15	(1.5)	0.13	(2.2)**
Responding parent very religious	-0.10	(1.6)	-0.23	(3.0)***	-0.12	(2.7)***
Father: high school graduate	-0.10	(1.4)	-0.10	(1.2)	-0.09	(1.6)
Father: >HS & <college grad	-0.07	(0.8)	-0.08	(0.8)	-0.20	(3.3)***
Father: $\geq$college graduate	-0.26	(2.6)***	-0.60	(4.8)***	-0.54	(7.5)***
Mother: high school graduate	0.05	(0.6)	-0.36	(4.0)***	-0.18	(3.4)***
Mother: >HS & <college grad	0.10	(1.2)	-0.32	(3.1)***	-0.37	(5.9)***
Mother: $\geq$college graduate	-0.03	(0.2)	-0.18	(1.4)	-0.61	(8.3)***
Household characteristics						
Number of younger siblings	0.03	(1.4)	0.04	(1.4)	0.06	(4.0)***
Number of older siblings	0.01	(0.2)	0.05	(1.8)*	0.03	(1.9)*
Father present in the household	-0.19	(3.1)***	-0.33	(4.5)***	-0.25	(5.7)***
Mother present in the household	-0.38	(3.9)***	-0.63	(5.8)***	-0.24	(3.3)***
Local environment characteristics						
Resides in suburbs	0.25	(3.7)***	0.18	(2.2)**	-0.12	(2.5)**
Resides in city	0.23	(3.3)***	0.28	(3.3)***	-0.05	(1.1)
Resides in North	0.14	(1.6)	0.37	(3.4)***	0.08	(1.3)
Resides in South	0.20	(2.6)**	0.20	(2.0)**	0.08	(1.4)
Resides in West	0.15	(1.8)*	0.19	(1.8)*	0.06	(1.0)
Unobserved heterogeneity						
Constant: u_1, v_1, ε_1	-1.56	(9.2)***	-2.83	(11.5)***	-5.54	(16.3)***
$u_{22}, v_{22}, \varepsilon_{22}$	-2.33	(14.3)***	-2.52	(20.6)***	-1.86	(14.5)***
α_1			1.88	(9.0)***		
α_2			1.49	(6.2)***		
α_3			-1.09	(3.0)***		
α_4			-0.50	(1.5)		
α_5			$-\infty$			
α_6			2.19	(10.3)***		
α_7			$-\infty$			
$-$Loglikelihood			27087.0			
(b) Independent unobserved heterogeneity						
Effect delinquency	–		–		0.28	(7.1)***
Effect arrest	–		–		0.52	(12.0)***
$-$Loglikelihood	10589.6		8126.7		8642.0	
LR test statistic			542.6**			

(c) Joint distribution of unobserved heterogeneity

		Arrest				
		High		Low		
		Delinquency				
School leaving	High	Low	High	Low	Total	
High	30	0	20	41	91	
Low	1	0	3	5	9	
Total (arrest)		31		69	100	

Absolute t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The structure of age dependence for school leaving: age dummies 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+. For arrest: age dummies 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+. For delinquency: 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+.

Table 5: Sensitivity

	Delinquency		Arrest		School Leaving	
(a) Baseline						
Effect delinquency	–		–		0.27	(5.5)***
Effect arrest	–		–		0.45	(9.0)***
−Loglikelihood			27087.0			
(b) Effect of leaving school at different ages on delinquency and arrest						
Effect delinquency	–		–		0.27	(5.6)***
Effect arrest	–		–		0.48	(9.5)***
Effect school leaving ≤ age 17	0.14	(0.6)	0.97	(7.6)***	–	
Effect school leaving at age 18 or 19	0.18	(1.0)	0.36	(3.1)***	–	
Effect school leaving ≥ age 20	-0.17	(0.4)	0.14	(0.9)	–	
−Loglikelihood			27058.8			
(c) Effect of delinquency on arrest						
Effect delinquency	–		0.25	(3.5)***	0.26	(5.4)***
Effect arrest	–		–		0.45	(9.0)***
−Loglikelihood			27082.5			
(d) Effect of initiating delinquency at different ages on arrest						
Effect delinquency	–		–		0.26	(5.4)***
Effect arrest	–		–		0.45	(9.0)***
Initiate delinquency ≤ age 17	–		0.29	(3.9)***	–	
Initiate delinquency ≥ age 18	–		0.06	(0.3)	–	
−Loglikelihood			27081.9			
(e) Effect of delinquency and arrest on school leaving by race						
Effect delinquency	–		–		0.40	(6.4)***
Effect delinquency × black	–		–		-0.28	(2.8)***
Effect delinquency × Hispanic	–		–		-0.34	(3.4)***
Effect arrest	–		–		0.44	(6.5)***
Effect arrest × black	–		–		0.08	(0.8)
Effect arrest × Hispanic	–		–		-0.04	(0.3)
−Loglikelihood			27079.0			
(f) Cross-sectional sample (N=3,393)						
Effect delinquency	–		–		0.32	(5.8)***
Effect arrest	–		–		0.47	(8.0)***
−Loglikelihood			20191.5			

Absolute t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the explanatory variables listed in Table 1. The structure of age dependence for school leaving: age dummies 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+. For arrest: age dummies 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+. For delinquency: 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+.

Table 6: Parameter estimates – multivariate hazard model: school leaving, arrest, income and non-income delinquency
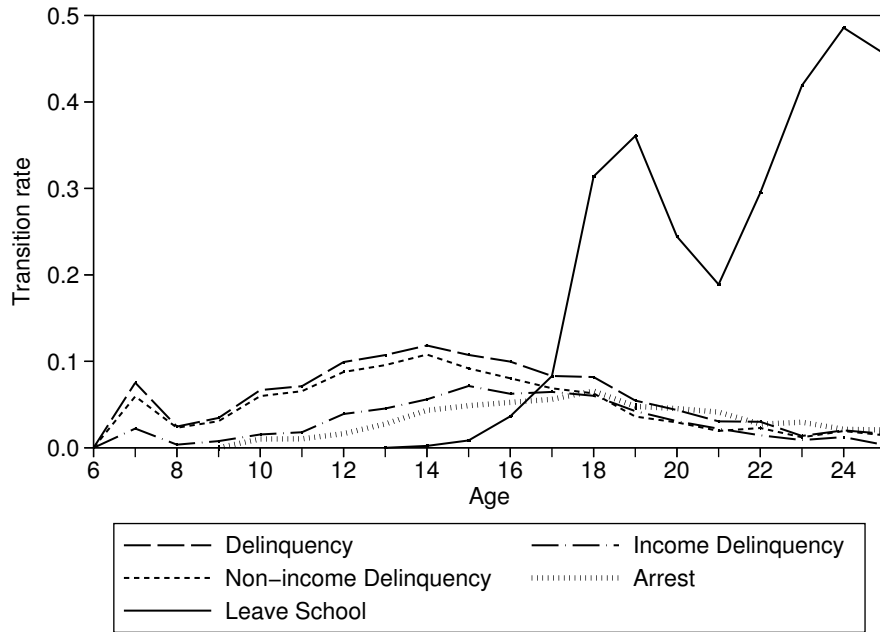
	Income Delinquency		Non-Income Delinquency		Arrest		School Leaving	
(a) Correlated unobserved heterogeneity								
Effect income delinquency	–		–		–		0.34	$(6.7)^{***}$
Effect non-income delinquency	–		–		–		0.09	$(1.8)^{*}$
Effect arrest	–		–		–		0.40	$(7.7)^{***}$
Unobserved heterogeneity								
Constant: $p_1, q_1, v_1, \varepsilon_1$	-2.54	$(11.5)^{***}$	-1.65	$(9.7)^{***}$	-2.70	$(11.3)^{***}$	-7.32	$(20.2)^{***}$
$p_{22}, q_{22}, v_{22}, \varepsilon_{22}$	-2.89	$(23.0)^{***}$	-2.62	$(16.7)^{***}$	-2.53	$(23.3)^{***}$	1.82	$(14.3)^{***}$
α_1					-3.46	$(9.7)^{***}$		
α_2					-3.18	$(10.8)^{***}$		
α_3								
α_4								
α_5					-0.67	$(9.9)^{***}$		
α_6					-1.45	$(13.0)^{***}$		
α_7					-3.57	$(9.9)^{***}$		
α_8								
α_9					-4.96	$(5.2)^{***}$		
α_{10}					-3.74	$(7.4)^{***}$		
α_{11}								
α_{12}					-2.16	$(10.4)^{***}$		
α_{13}					-2.60	$(10.8)^{***}$		
α_{14}					-1.39	$(8.6)^{***}$		
α_{15}								
$-$Loglikelihood				33913.7				

	Income Delinquency		Non-Income Delinquency		Arrest		School Leaving	
(b) Independent unobserved heterogeneity								
Effect income delinquency	–		–		–		0.34	$(7.9)^{***}$
Effect non-income delinquency	–		–		–		0.10	$(2.3)^{**}$
Effect arrest	–		–		–		0.47	$(10.6)^{***}$
$-$Loglikelihood	7972.5		10026.0		8126.7		8622.7	
LR test statistic				1668.4**				

	Income Delinquency		Non-Income Delinquency		Arrest		School Leaving	
(c) Effects of first delinquency and first arrest at different ages on school leaving								
Effect income delinquency								
Initiate income delinquency $\leq$ age 15	–		–		–		0.37	$(6.4)^{***}$
Initiate income delinquency at age 16 or 17	–		–		–		0.29	$(3.6)^{***}$
Initiate income delinquency $\geq$ age 18	–		–		–		0.32	$(3.0)^{***}$
Effect non-income delinquency								
Initiate non-income delinquency $\leq$ age 15	–		–		–		0.08	(1.4)
Initiate non-income delinquency at age 16 or 17	–		–		–		0.01	(0.1)
Initiate non-income delinquency $\geq$ age 18	–		–		–		0.27	$(2.0)^{**}$
Effect arrest								
First arrest $\leq$ age 15	–		–		–		0.54	$(8.8)^{***}$
First arrest at age 16 or 17	–		–		–		0.32	$(4.3)^{***}$
First arrest $\geq$ age 18	–		–		–		0.14	(1.4)
$-$Loglikelihood				33903.2				
LR test statistic				21.0				

(d) Joint distribution of unobserved heterogeneity

	Arrest							
	High				Low			
	Income delinquency				Income delinquency			
	High		Low		High		Low	
	Non-income delinquency				Non-income delinquency			
School leaving	High	Low	High	Low	High	Low	High	Low
High	22	1	3	0	10	0	11	43
Low	2	0	0	0	2	0	1	5
Total (arrest)	28				72			

Absolute t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The structure of age dependence for school leaving: age dummies 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+. For arrest: age dummies 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+. For income and non-income delinquency: 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25+.

Table 7: Effect on cumulative probability of school leaving for a reference individual who is susceptible to early school leaving; simulations (%)

Age	(1) No delinquency; no arrest	(2) Delinquent at 16; no arrest	(3) Delinquent at 16; arrest at 17	(4) Delinquent at 14; no arrest	(5) Delinquent at 14; arrest at 15
14	0	0	0	0	0
15	1	1	1	1	1
16	4	4	4	5	7
17	10	12	12	13	19
18	35	42	55	43	58
19	60	69	83	69	84
20	73	81	92	82	93
21	81	88	96	89	97
25	100	100	100	100	100

Figure 1: Transition rates for first delinquency, arrest and school leaving by age

Figure 2: Cumulative starting probabilities for first delinquency, arrest and school leaving by age
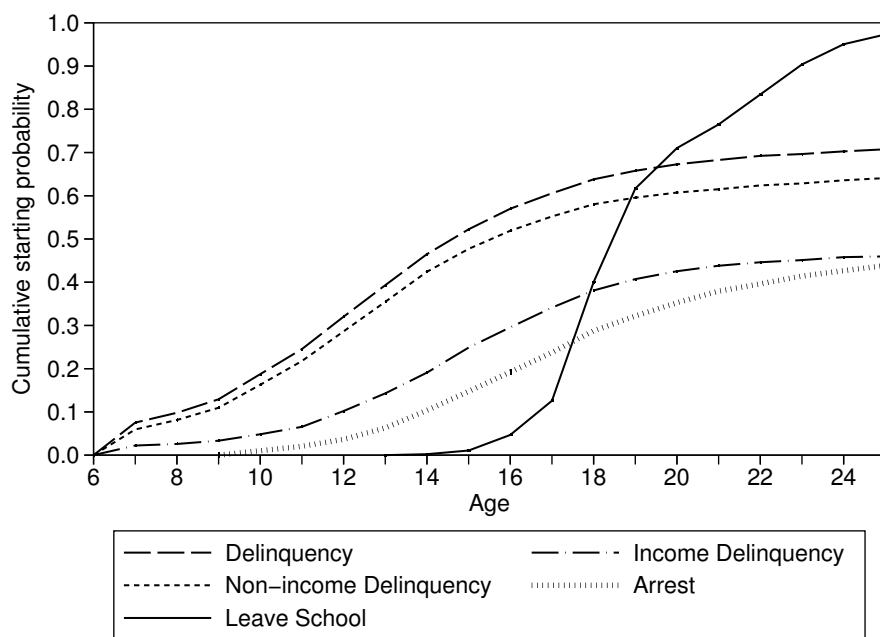
Table A.1: Age pattern

Age	Left School		Committed Income Delinquency		Committed Non-income Delinquency		Committed Delinquency		Arrested	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
7	0	0	0	100	0	267	0	338	0	0
8	0	0	0	16	0	98	0	102	0	0
9	0	0	0	34	0	127	0	140	0	0
10	0	0	0	67	0	239	0	261	0	46
11	0	0	0	77	0	246	0	260	0	46
12	0	0	5	164	4	308	3	336	2	72
13	0	0	7	183	9	306	8	327	4	119
14	26	10	27	215	24	311	27	321	15	182
15	21	38	39	257	26	234	25	254	20	195
16	26	160	29	206	28	184	24	208	23	199
17	38	349	55	198	43	143	38	154	31	200
18	26	1,199	70	169	57	119	44	136	28	219
19	18	936	346	109	225	62	210	81	32	148
20	12	401	375	65	261	42	223	52	37	132
21	13	232	324	37	230	22	191	28	33	113
22	13	290	290	19	197	20	167	21	55	71
23	8	285	271	9	184	8	155	7	57	73
24	10	188	74	9	38	9	30	7	110	50
25	19	86	147	2	77	6	63	5	467	45
26	13	39	160	3	93	7	78	4	459	26
27	6	14	117	2	85	0	54	0	439	15
28	3	5	117	0	76	1	52	1	404	6
29	1	3	90	3	70	1	49	3	309	3
30	0	0	1	0	1	0	1	0	3	0
Total	253	4,235	2,544	1,944	1,728	2,760	1,442	3,046	2,528	1,960
	4,488		4,488		4,488		4,488		4,488	