I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14129

High School Dropout and the Intergenerational Transmission of Crime

Davide Dragone
Giuseppe Migali
Eugenio Zucchelli

FEBRUARY 2021

DISCUSSION PAPER SERIES

IZA DP No. 14129

High School Dropout and the Intergenerational Transmission of Crime

Davide Dragone
University of Bologna

Giuseppe Migali
Lancaster University Management School and Magna Graecia University

Eugenio Zucchelli
MIAS, Universidad Autónoma de Madrid, Lancaster University and IZA

FEBRUARY 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

High School Dropout and the Intergenerational Transmission of Crime

We explore the relationship between high school dropout and pupils' adult crime by accounting for the role of the intergenerational transmission of crime. We employ a human capital model of schooling and crime and show that the intergenerational transmission of crime could have a direct effect on adult crime as well as an indirect effect mediated by high school dropout. We empirically assess the relevance of these relationships using fixed effects linear probability models and inverse probability weighting regression adjustment on US data from the National Longitudinal Study of Adolescent to Adult Health. We find that dropping out from high school and having a convicted father increase the probability of adult crime, with the former presenting a larger effect. Our empirical models also suggest that having a convicted father increases the probability of dropping out from school. This reveals that paternal crime imposes a double penalty on children: it increases their probability of committing crimes later on in life both directly and indirectly via school dropout. When considering the role of the environment, we find that while an early exposure to high levels of crime exacerbates dropping out, it has no direct long-term effect on adult crime. Finally, we show that individual traits may also play a role, as pupils with lower levels of cognitive skills present higher probabilities of adult criminal behaviour and stronger intergenerational effects.

JEL Classification: I26, J62, K42

Keywords: high school dropout, crime, intergenerational transmission, Add Health

Corresponding author:
Eugenio Zucchelli
Department of Economic Analysis
Economic Theory and Economic History
Universidad Autónoma de Madrid (UAM),
C/Francisco Tomás y Valiente 5
28049, Madrid
Spain

E-mail: eugenio.zucchelli@uam.es

1 Introduction

Dropping out from high school and crime are two major and closely interrelated policy issues. Early school attrition is a widespread phenomenon, concerning around 15% of pupils in OECD countries (OECD, 2019), with serious long-term consequences. These include an increased risk of unemployment; lower lifetime earnings; poor health; as well as a higher propensity to criminal behaviour (e.g. Cook and Kang, 2016; McFarland et al., 2018). Crime itself is systematically linked to poverty, social exclusion and reduced employment opportunities (Freeman, 1999). As a result, school dropout and crime impose substantial economic costs: for instance, in the US alone the combined income and tax losses for one cohort of 18 years old of dropouts is around 1.6% of the GDP, whereas public expenditure on criminal justice alone absorbs nearly 2.5% of the GDP (Chalfin, 2015; Monrad, 2007). Hence, strategies to increase school completion and reduce crime feature prominently in the policy agenda of any government.

While the association between education and crime is well-documented and there is increasing evidence supporting a causal link (Machin et al., 2011; Cook and Kang, 2016), less is known on the mechanisms governing such relationship. Standard economic models of crime suggest that higher levels of human capital may reduce crime if they increase the expected return from a legitimate job versus the one of a criminal activity (Freeman, 1999). Yet, such models do not typically account for the effects of the intergenerational transmission of crime. In addition, previous empirical papers have separately analysed either the relationship between education and crime or the intergenerational transmission of crime. A more accurate knowledge of the different pathways leading dropouts to committing crimes would allow policy makers devising targeted strategies to curb criminal behaviour.

The main objective of this paper is to investigate the relationship between high school dropout and pupils' adult crime while considering the specific role played by the intergenerational transmission of crime. We propose a two-period human capital model of crime to describe the trade-off between legitimate versus criminal activities (Becker, 1968; Freeman, 1999; Lochner, 2004). Our model accounts for the effects of fathers' criminal behaviour on both dropping out from school and committing crimes in adulthood. It also allows for the exploration of wider mechanisms affecting adult crime such as the early exposure to different levels of crime, cognitive skills, and parental nurture. We first estimate a series of fixed effects linear probability models to establish baseline correlations among our variables of interest. We then employ inverse probability weighting regression adjustment (IPWRA) models to more precisely identify direct and indirect effects of the intergenerational transmission of crime. IPWRA models involve a two-stage estimation which enables analysing the direct effect of a father's conviction on adult crime and its indirect effect via high school dropout. We exploit uniquely suited data drawn from the National Longitudinal Study of Adolescent to Adult Health (Add Health) providing detailed information on students' education and criminal behaviour, fathers' convictions together with a wealth of further individual- and family-level variables.

Our paper builds on and merges two strands of the literature. A first strand includes studies on both contemporaneous and long-term effects of education on crime in different disciplines. Early works generally find that dropping out from school is positively associated with juvenile criminal behaviour (Thornberry et al., 1985; Fagan and Pabon, 1990), while subsequent studies provide mixed results on the effects of additional years of education or having earned a high school degree on juvenile or later crime (Tauchen et al., 1994; Grogger, 1997; Lochner, 2004; Buonanno and Leonida, 2009). Quasi-experimental studies exploiting changes in the mandatory school leaving age in several countries consistently suggest the presence of a causal effect of education on crime (Lochner and Moretti, 2004; Machin et al., 2011; Anderson, 2014; Hjalmarsson et al., 2015). A separate line of research looks at the intergenerational transmission of crime and generally finds it to be persistent across generations. For instance, Duncan et al. (2005) find that children with convicted mothers are around five times more likely to commit juvenile crimes, and results from Hjalmarsson and Lindquist (2012) indicate that individuals with sentenced fathers have between 2.06-2.66 higher odds of receiving a criminal conviction. However, and to the best of our knowledge, none of these earlier studies have jointly accounted for the effects of both school dropout and the intergenerational transmission of crime on adult criminal behaviour, explored their interactions and discussed specific mechanisms linking them.

Our findings suggest that both dropping out from high school and having a convicted father increase the probability of adult crime, with the former playing a larger role. Importantly, we find that paternal crime could have direct as well as indirect effects on pupil's adult crime, through high school dropout. That is, a father's conviction raises the likelihood of dropping out, which in turn increases the probability of committing crimes when adult. While an early exposure to high crime levels does not have a long-term effect on adult crime, it appears to have a significant short-term influence on early school attrition. Lower levels of cognitive skills are systematically linked to higher probabilities of crime later in life and stronger intergenerational effects. A father's nurture also appears to be relevant: the absence of a (potentially convicted) biological father seems to halve the persistence of crime between generations. These results are robust to alternative definitions of crime and the inclusion of personality traits.

This paper contributes to the literature in several important ways. First, our approach allows for a more comprehensive examination of different pathways leading to adult crime. It does so by acknowledging and testing the presence of direct and indirect effects of the intergenerational transmission of crime in the relationship between high school dropout and adult crime. Indeed, we provide evidence of significant direct and indirect effects of paternal crime and ignoring them could produce different findings and policy recommendations. This contributes directly to the literature on schooling and crime which does not often account for the intergenerational effects of criminal behaviour. Second, our data permits exploring concurrent and well-established mechanisms affecting dropping out from school and adult crime such as the early exposure to different levels of crime and the influence of cognitive skills and nurture. Third, this paper also contributes to the literature on the intergenerational transmission of crime. Our findings confirm the strength of the association of criminal behaviour across generations, and they support the idea that high school completion is an important tool to mitigate the impact of a father's criminal behaviour. Overall, we advocate that studies aimed at investigating the relationship between schooling and crime should consider the intergenerational effects of crime.

2 A human capital model of crime

In order to explore the mechanisms linking high school dropout, pupils' adult crime and the intergenerational transmission of crime, consider a schooling and labour supply problem over two time-periods. In period 1, the individual is a teenager choosing how to allocate his time endowment T_1 between schooling related activities s and non-schooling related activities n. Following Becker (1994), schooling time increases human capital according to

$$H = h\left(s; X\right), \tag{1}$$

where $\partial H/\partial s > 0$. $\partial^2 H/\partial s^2 < 0$, and $X = \left(X^1, X^2, ..., X^f, ..., X^K\right)$ represents a vector of exogenous individual and household characteristics such as individual skills, socioeconomic status and household income; among these variables, X^f denotes whether the pupil's father was convicted.

The teenager's non-schooling time contributes to period-1 income (through, e.g. part-time jobs), according to

$$M_1 = m\left(n; X\right) \tag{2}$$

Income M_1 (which also depends on family income, a component of X) becomes consumption y_1, that the teenager enjoys according to an increasing and concave per-period utility function $U_1\left(y_1; X\right)$. Since $\partial M_1/\partial n > 0$, non-schooling time increases period-1 income and, due to the time constraint $T_1 = s + n$, it also reduces the schooling time needed to increase human capital.

In period 2 the individual becomes an adult and divides his time endowment T_2 between legitimate jobs l and criminal (non-legitimate) activities c. Both activities contribute to period-2 income, depending on human capital H and the vector of exogenous variables X, according to[1]

$$M_2 = m_l\left(l; H, X\right) + m_c\left(c; H, X\right) \tag{3}$$

Note that illegal income $m_c\left(c; H, X\right)$ is in expected terms as it depends on the probability of the corresponding criminal activity being sanctioned by the authority. We assume that expected income for each activity is increasing in the time allocated to it, and that income is increasing in human capital, with decreasing returns. Period-2 consumption y_2 equals M_2 and yields utility $U_2\left(y_2; X\right)$.

2.1 Solving the model

Denoting with β the time discounting factor, the intertemporal problem of the agent is to maximise

$$U_1\left(y_1; X\right) + \beta U_2\left(y_2; X\right) \tag{4}$$

subject to the per-period budget constraints $y_1 = M_1$ and $y_2 = M_2$, the time constraints $T_1 = s+n$ and $T_2 = l + c$, and the process of human capital accumulation $H = h\left(s; X\right)$. The model is solved by backward induction.

In period 2, criminal behaviour depends on the gap Δ between the marginal income obtained by investing time on criminal activities, and the marginal income obtained through legitimate activities:[2]

$$\Delta \equiv \frac{\partial m_c(c; H, X)}{\partial c} - \frac{\partial m_l(l; H, X)}{\partial l} \tag{5}$$

Three possible cases can arise: I. when the crime-labour marginal income gap is zero, the agent divides his working time between legitimate and criminal activities; II. if $\Delta > 0$, he exclusively engages in criminal activities; III. if $\Delta < 0$, it is optimal for the agent to only engage in legitimate jobs. Since the marginal income of labour and crime depends on the human capital acquired through schooling time when young, as well as the exogenous variables represented by X, we denote the optimal time allocated to crime as $c^* = c(H, X)$.

In period 1, the teenager chooses the amount of time s^* allocated to schooling vs non-schooling related activities. On one hand, schooling decreases period-1 utility as it reduces time for leisure as well as time that could be otherwise employed to produce income. On the other hand, schooling is an investment which pays off when adult. Specifically, it increases human capital, which in turn will increase the marginal productivity of work, income and consumption. Considering an internal

[1] To simplify the exposition, we assume that working time only indirectly affects utility through income and has no direct effect on utility. This assumption can be relaxed to account for, e.g., the effort cost of work or the disutility of being arrested and convicted, however this would just complicate the notation without providing any further insight.

[2] All proofs are in the Appendix.

solution, optimal schooling time s^* satisfies (henceforth omitting the arguments for brevity)

$$\frac{\partial U_1}{\partial y_1}\frac{\partial m}{\partial n} = \beta \frac{\partial U_2}{\partial y_2}\frac{\partial y_2}{\partial H}\frac{\partial H}{\partial s} \tag{6}$$

where the left-hand side represents the marginal period-1 utility from non-schooling related activities, and the right-hand side describes the impact of schooling on future income (either from labour or crime) and future discounted utility.

Note that a necessary condition for some schooling to be chosen is that it increases the expected income in period 2, i.e. $\partial y_2/\partial H > 0$. This is typically more relevant for skilled work (requiring high human capital) rather than for unskilled work. Accordingly, it is likely that a teenager will drop out from high school if he/she anticipates that education will not contribute enough to his/her future income. Such assessment depends on both current and future exogenous variables, as well as the degree of the individual's impatience, as represented by the time discounting factor β.

2.2 Intergenerational transmission of crime and other mechanisms

We are interested in investigating how criminal activities in period 2 are influenced by education and other period-1 variables described by vector X, with a focus on the role of paternal crime. More specifically, we first consider the effect of human capital on crime dc^*/dH, and then study the roles of key exogenous variables in influencing schooling and crime.

Concerning education, it is apparent that human capital will decrease the time allocated to criminal activities if it increases the marginal productivity of a legitimate job versus the one of an illegitimate (criminal) activity. Formally, this is described by

$$\frac{dc^*}{dH} = \xi \frac{\partial \Delta}{\partial H} \tag{7}$$

where ξ is positive by concavity of 3 and the term $\partial\Delta/\partial H$ denotes the effect of H on the crime-labour marginal income gap. Under the assumption that education makes legitimate activities relatively more lucrative than criminal ones, the sign of the gap would be negative. This implies that pupils who drop out from school are expected to be more likely to engage in criminal activities when adults, with respect to those who complete high school. We test this assumption in our empirical analysis.

In addition, we explore the potential intergenerational transmission of crime from fathers to their children. As explained in the empirical section, this will be assessed by estimating the relationship between the pupil's adult crime and whether the pupil's father was convicted when he or she was at school. Recall that X^f denotes whether the father was convicted. The intergenerational transmission of crime can be then decomposed as follows[3]

$$\frac{dc^*}{dX^f} = \xi \frac{\partial \Delta}{\partial X^f} + \frac{\partial H}{\partial X^f}\frac{dc^*}{dH} \tag{8}$$

The first term represents the direct channel through which the intergenerational transmission of crime occurs. This passes through the effect of having a convicted father on the labour-crime marginal income gap. Importantly, this first direct effect can be due to shared criminal opportunities and socioeconomic conditions, nurture, family or social norms. The second term describes the effect of a convicted father on adult crime *through* the pupil's choice of going to school versus dropping out. This indirect channel thus highlights the possible mediating effect of education, which can be affected by the presence of a convicted father (as described by $\partial H/\partial X^f$). This will ultimately have an impact on adult crime ($\partial c^*/\partial H$) by also affecting the labour-crime marginal income gap (see eq. 7).

We will explore such effects using our empirical models. Specifically, we will first use linear probability models to establish baseline correlations between high school dropout and pupils' adult crime, while accounting for fathers' convictions and other period-1 variables. We will then use inverse probability weighting regression adjustment (IPWRA) models to estimate the effects of X^f on high school dropout (which provides information on $\partial H/\partial X^f$) and the ones of early school attrition and X^f on adult crime (i.e. $\partial c^*/\partial H$ and $\partial c^*/\partial X^f$). These should shed light on the direct effects of fathers' convictions on pupils' adult crime versus indirect effects via high school dropout.

Equation 8 is sufficiently general to also describe the effect of any exogenous variable $X^k \in X$, $k \in \{1, 2, ..., K\}$. Accordingly, it provides an opportunity to dig deeper into the mechanisms linking dropping out from school to adult crime. In our empirical analysis, we pay particular attention to three variables whose relevance is reported in the literature. First, we account for the effects of early exposure to a broad criminal environment on pupils' adult crime, using information on the number of crimes committed in the county where individuals resided when going to school. This is supported by quasi-experimental evidence on the effect of early exposure to neighbourhood crime on subsequent criminal behaviour (Damm and Dustmann, 2014), as well as on the effects of place of residence during childhood on earnings and education patterns (Chetty and Hendren, 2018). Second, we explore the role of cognitive skills, given their well-know interplay with educational outcomes and a number of social behaviours throughout the life-cycle, including crime (e.g. Cunha et al., 2006). Third, we deal with the standard nurture vs nature issue in intergenerational studies (Black et al., 2011), by exploiting information on the absence of a biological father within

[3]Since X^f is a discrete variable, it would be more appropriate to consider a discrete change rather than a derivative. However, here we stretch the notation to provide a result that also allows describing the effect of marginal changes of a generic variable $X^k \in X$.

the pupils' families. This allows better isolating the effect of a (potentially convicted) father's nurturing on adult crime. In our empirical models, all these variables are considered together with a wide range of further demographic and socioeconomic individual-level and family-level observables also included in X.

3 Data

Our empirical analysis is based on detailed information on education and crime drawn from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a panel study of a representative sample of US high school students initially in grades 7-12 and subsequently followed through adolescence and transition into adulthood. For the purpose of our study we employ the first four waves of Add Health, that is until individuals in the sample are aged 24-32 (wave I: 1994; wave II: 1995-1996; wave III: 2001; wave IV: 2008).[4]

Add Health is uniquely suited to our empirical approach as it includes a wide range of variables on the main respondents' socioeconomic status; psychological and physical well-being, including several measures of physical and mental health; cognitive and noncognitive traits; risky attitudes; health behaviours as well as pupils' and fathers' criminal behaviour. It also collects data on other parents' characteristics, neighbourhood, schools as well as county-level contextual data.

Add Health consists of four main questionnaires: the in-home questionnaire; the parent questionnaire; the school questionnaire and the school-administrator questionnaire. The core sample of the wave I in-home questionnaire includes 12,105 students randomly selected from the 132 schools. Booster samples based on ethnicity (Cuban, Puerto Rican and Chinese), adoption and disability status were also added to the core sample together with an additional sample of African-American students with highly educated parents, leading to a total of 20,745 individuals. This augmented sample provides the basis for the four longitudinal follow ups we employ. Importantly, the parent questionnaire provides data on marriage, health, education and employment and should be completed by all parents of those pupils responding to the wave I in-home questionnaire. The in-school and school-administrator questionnaires were collected between 1994-1995 and include further information on school activities, friendship networks and a series of health conditions as well as further school context data and policies, respectively.

Given the main objective of this paper, we focus on pupils initially enrolled at school. Although we combine data from wave I with multiple retrospective information on degrees completion collected in waves II to IV, this essentially corresponds to all the adolescents interviewed in wave I. We draw information on pupils' individual-level characteristics, covering education and crime, as

[4] Add Health sample has a school-based design including 132 schools stratified by region, urban area, size, school type and ethnicity to ensure representativeness among US schools. For further details on the sampling strategy, see Harris et al. (2009).

well as fathers' crime from the in-home questionnaire. We then merge this with further parental background information from the parent questionnaire; school-level characteristics (school size and type, alternatively replaced with school fixed effects) from the school questionnaires; and county-level information on the total number of crimes included in wave I contextual data files. [5]

3.1 Key variables

Our key variables include pupils' adult crime, our main outcome of interest, as well as high school dropout and fathers' criminal behaviour. Pupils' adult crime is defined as a dummy variable taking value 1 if individuals were ever arrested for any type of offence after 18 years old, 0 otherwise. As a robustness check, we also employ an alternative and more restrictive definition of adult crime based on whether pupils were convicted and spent time in jail (at least once after 18 years old), 0 otherwise. Information on pupils' adult crime is based on a series of questions about ever been arrested, age of first arrest and convictions/jail time asked in wave IV.

High school dropout is a binary variable which equals 1 if pupils dropped out from school, 0 if they have at least a high school degree. This definition of school dropout includes adolescents reporting "dropout" or "other non-graduate" in their high school exit status in wave III. This information is then cross-checked with the highest education level reported in wave IV, i.e. "8th grade or less", "some high school", or "did not earn diploma". Initially, we also include General Education Development (GED) certificate holders among dropouts as these individuals must first drop out before earning a degree equivalent to a high school diploma. However, we also estimate our models by excluding GED holders and thus restricting our sample to"pure" dropouts. This further analysis is performed in order to test the role of cognitive skills in influencing adult crime as pure dropouts appear to systematically present lower cognitive skills if compared to GEDs (Heckman and LaFontaine, 2010).

Fathers' criminal behaviour is also defined as a binary variable which equals 1 if the biological father spent time in jail when the pupil was less than 18 years old and 0 otherwise. More specifically, we use information from the following wave IV questions: "How old were you when your biological father went to jail/prison (for the first time)?" and "Has/did your biological father ever spent some time in jail/prison?". We focus on father's criminal behaviour in the first 18 years of the pupil's life to avoid potential simultaneity issues with the variable defining pupils' adult crime.

[5] In this paper, we show estimates produced with models including school fixed effects. Estimates produced by corresponding models including school characteristics are very similar and available upon request.

3.2 Other variables

Our empirical models also account for a wide range of further individual- and parental-level characteristics as well as broader environmental conditions. Among the individual-level variables, we include: gender; ethnicity (white, hispanic, African-American, Asian); physical and mental health (a binary measure based on self-reported ill-health for the former and a second binary indicator identifying individuals feeling depressed all time/most of the time for the latter); a well-established general measure of cognitive skills (the abridged version of the Peabody Picture Vocabulary Test, PPVT); and learning disabilities (a dummy variable for the presence of difficulties with attention, dyslexia, or other reading, spelling, writing, or math disability).

Following (Young and Beaujean, 2011), we exploit information drawn from 13 questions included in both the in-home and in-school questionnaire of wave I to build three of the so called Big Five personality traits: conscientiousness, neuroticism and extroversion. These traits should account for a series of personality facets such as reliability and self-discipline; emotional stability; as well as proneness towards positive emotions and sociability, respectively, that could be relevant for our analysis and that have been previously linked to both criminal behaviour (O'Riordan and O'Connell, 2014) and school dropout (Migali and Zucchelli, 2017).

We also include a series of variables accounting for an individual's attitude towards risk and other risky behaviours beyond crime. These include a general risky attitude binary indicator capturing at least one of the following behaviours: no use of seat belts or no use of birth controls, and a binary variable for health-behaviours (that equals 1 in the presence of the frequent consumption of at least one the following: tobacco, alcohol, marijuana or other drugs; 0 otherwise).[6] In addition, we account for time preferences (myopic behaviour) via a dummy variable which equals 1 when an individual reports to agree/strongly agree to the sentence "I live my life without much thought for the future"; 0 otherwise. Finally, since religious affiliation has been linked to a lower likelihood to commit crimes (e.g. Iyer, 2016), we include a dummy variable for whether pupils follow any religious denominations.

Broader parental- and family-level variables comprise education levels of mother and fathers (primary, secondary and higher education vs no education as baseline category) and their employment (routine/technical job; small employer/intermediate job; managerial/professional job vs unemployed/home-maker as baseline category). We also account for risky health-behaviours of family members by employing a binary indicator capturing at least one of the following behaviours: main parent is a smoker; another member of the family is a smoker; main parent drinks

[6]More specifically, this variable captures at least one the following behaviours: smoking 20 or more cigarettes per day (heavy smoking); having 5 or more drinks each time you drink (heavy alcohol consumption); having 30 or more marijuana cigarettes in the last 30 days (daily marijuana consumption); using cocaine or inhalants 10 or more times in the last 30 days (frequent use of other drugs). Since our main objective is to account for other variables potentially influencing criminal behaviour as well as school dropout, we chose to focus on frequent or heavy consumption of the above mentioned drugs. More broadly, these could also reflect time and risk preferences.

more than 5 five drinks at times at least 3 times a week. In order to test the influence of parents' nurturing, we also account for whether pupils lived with or without their biological fathers or mothers when at high school.

3.3 Descriptive statistics

Table 1 reports basic descriptive statistics for our estimating sample. According to our initial definition of pupils' adult crime, 22.7 percent of pupils are arrested at least once when adult (i.e. 18 years old or over) for any type of offence. This percentage decreases to 15.6 when considering a stricter definition of adult crime based on being convicted at least once. As for the other key variables, around 9.3 percent of fathers spent time in jail and about 7.8 percent of pupils appear to have dropped out from high school. Nearly 45 percent of individuals in our sample are male; around 53 percent are white, while 14.3 percent identify as Hispanic, nearly 22 percent as African-American and about 8 percent as Asian. Only a relatively small proportion of pupils report being in ill-health (6.6 percent) while a slightly higher percentage report being depressed (9.8 percent).

The Peabody Picture Vocabulary Test (PPVT), a reliable and widely used measure of basic cognitive skills (Dunn and Dunn, 2007), presents a mean value of about 101, a score considered to be within the expected average values. Furthermore, nearly 9.5 percent of individuals are affected by a learning disability.

With respect to noncognitive skills, 14.4 percent of pupils report low levels of conscientiousness; 7.7 percent are classified as neurotic; and nearly 12 percent are found to be introverted. In terms of behaviours implying some form of risk, around 33.5 percent of pupils report having a general risky attitude (as defined by no use of seat belts or birth controls); 12.5 percent are myopic/impatient; and 22.6 percent present at least one risky health-behaviour. Around 87 percent of individuals report following a religious denomination.

4 Empirical approach

Our empirical strategy exploits the wide range of observed variables included in Add Health, covering detailed information on education as well as criminal behaviour of both pupils and their fathers. We first estimate a series of linear probabilities models (LPM) with state and school fixed effects to establish correlations between our variables of interest. We further employ inverse probability weighting regression-adjustment (IPWRA), a quasi-experimental approach capable of providing measurements of unobserved potential outcomes (Imbens and Wooldridge, 2009; Cattaneo, 2010). In addition, IPWRA involves a two-stage estimation process reflecting key elements of our human capital model of crime. More specifically, the first stage estimates selection into treatment (high school dropout), conditional on observables, and computes inverse probability

weights. This stage accounts for the effects of (period-1) observed variables on adult crime, as described by dc^*/dX^k in equation 8 (or the corresponding expression for dc^*/dX^k). The second stage employs the estimated inverse probability weights to fit weighted regression models of the outcome (pupils' adult crime) in the presence and absence of high school dropout, and computes corresponding means of the treatment-specific predicted outcomes. The difference of such predicted outcomes provides an estimate of the average treatment effect. Hence, the second stage sheds light on direct and indirect channels affecting adult crime; the intergenerational transmission of crime; and the roles of other observed variables.
Accordingly, we first estimate a series of LPM based on the following specification:

$$E[Y_i|X_i] = \gamma + \mu_s + \theta_c + \alpha D_i + \delta FC_i + \eta \mathbf{X_i}. \tag{9}$$

where the information contained in X_i is decomposed into school dropout (D_i); a father's conviction (FC_i); and broader individual and family characteristics ($\mathbf{X_i}$). Specifically: D_i is a dummy variable taking value 1 if students dropped out from high school (0 if they have at least a high school degree) and FC_i is a binary variable representing whether their biological fathers spent time in jail (when the pupils were less than 18 years old), 0 otherwise. μ_s are school fixed effects accounting for any school-specific time-invariant factors and θ_c are pseudo-state fixed effects, controlling for any systematic variation across US states. Parameter α captures the effect of our main variable of interest (high school drop out) on adult crime, while parameter δ describes whether the pupil's father was convicted. Hence, the latter allows us accounting for the intergenerational correlation between fathers' crime and children's adult crime. Vector $\mathbf{X_i}$ includes the set (or subset, depending on the specific model) of observed exogenous variables describing further individual and family characteristics.

Secondly, we employ IPWRA. In our case, the first stage (treatment model) simple translates into the estimation of a probit model for our treatment (dropping out from high school) including the full set of covariates:

$$p(X) = P(D_i = 1|FC_i, \mathbf{X_i}) = \Phi(\gamma + \delta FC_i + \eta \mathbf{X_i}) \tag{10}$$

where Φ is the cumulative standard normal distribution function and D_i is the corresponding treatment (dummy) variable.

Using this model, we can predict the conditional probability of dropping out from school for each observation in the data, $\widehat{p}(\mathbf{X_i})$. We then generate inverse probability weights so that each individual's weight is equal to the inverse of the probability of the actual treatment received. Precisely,

$$w_i = \frac{1}{\widehat{p}(\mathbf{X_i})} \quad \text{if} \quad D_i = 1$$

are used to weigh observations in the treatment group, so that weights will be large when the probability of dropping out is small. Whereas

$$w_i = \frac{1}{1 - \widehat{p}(\mathbf{X_i})} \quad \text{if} \quad D_i = 0$$

weigh observations in the control group, so that weights will be large when the probability of obtaining at least a high school degree (not dropping out) is small. This ultimately translates into assigning larger weights to treated individuals (dropouts) with similar observables to untreated individuals (non dropouts), and therefore creating counterfactuals that are as comparable as possible to treated individuals.

The second stage consists of the estimation of separate linear weighted regression models for the two values of the treatment (outcome models), that is alternatively for dropouts and non dropouts. Importantly, observations in these models are weighted using inverse probability weights w_i, generated from the predicted conditional probabilities obtained in the first stage.

The difference between the potential outcomes provides the estimate of the average treatment effect of high school dropout on adult crime:

$$\tau_{ATE} = N^{-1} \sum_{i=1}^{N} (E[Y_i|FC_i, \mathbf{X_i}, D_i = 1] - E[Y_i|FC_i, \mathbf{X_i}, D_i = 0]) \tag{11}$$

Note that this is a consistent estimator if standard assumptions around conditional independence, $(Y_1, Y_0 \perp\!\!\!\perp D|p(X))$, and common support, $(0 < Pr(D = 1|X) < 1)$, hold. Finally, IPWRA estimators have an attractive doubly-robust property, implying that estimates will be consistent even if one of the two models is not correctly specified (Cattaneo, 2010).

5 Main results

Table 2 presents key estimates obtained using a series of incrementally comprehensive linear probability models (LPM). These allow investigating correlations between high school dropout, fathers' and children's criminal behaviour. More specifically, columns 1-3 and columns 4-6 of Table 2 focus on the effects of dropping out from high school on adult crime and the intergenerational transmission of crime, respectively. Column 7 reports estimates of the most comprehensive linear probability model including the effects of both high school dropout and fathers' convictions on pupils' adult crime. Looking at columns 1-3, we notice that the correlation between early school attrition and adult crime appears to be highly statistically significant throughout all specifications. In the absence of any other observed characteristics as well as fixed effects, school attrition is associated with an increased probability of committing a crime when adult of around 23.8 percentage points (pp) (column 1). This correlation decreases to 22.8pp when including state and school fixed effects (column 2) and to 18.1pp when also adding the full set of individual- and family-level observed characteristics (column 3). Similarly, in columns 4-6 we observe that the estimated coefficient for the intergenerational transmission of crime ranges between 11.5-15.3pp, depending on the model, and it is always highly statistically significant. Importantly, the most comprehensive model shows highly statistically significant coefficients for both high school dropout and fathers' convictions with corresponding effects of around 16.4 and 10.8pp. Overall, this initial set of LPM estimates suggests that, conditional on a wide range of observables and different types of fixed effects, the correlation between high school dropout and adult crime is larger than that estimated for the intergenerational transmission of crime.

Results reported in Table 3 further explore the complex relationships between high school dropout, pupils' adult crime and paternal convictions, employing inverse probability weighting regression adjustment (IPWRA). More specifically, column 1 reports results of the first stage treatment model of our baseline IPWRA specification, exploring the determinants of high school dropout. Accordingly, we find that among a wide range of determinants a father's conviction greatly increases the probability of dropping out from high school (64.4pp).[7] Columns 2 and 3 show results from the outcome models for adult crime, i.e. the estimated effects of a father's conviction on adult crime for pupils who did not and did dropout from school, respectively. Such models employ inverse probability weights produced during the first stage of the IPWRA estimation and include the full set of controls. While the intergenerational effect of crime is highly statistically significant in both models, the corresponding estimated coefficient is larger for pupils dropping out from school (13.7pp vs 11.6pp). This appears to imply that dropping out from high school may strengthen the effect of a father's conviction on pupils' adult crime. Finally, Table 3 also shows the average treatment effect (ATE) of dropping out from high school on adult crime. This is computed as the the difference between the predicted outcomes of the second stage weighted regression models. The ATE is highly statistically significant and similar in size to the corresponding correlation obtained using the most comprehensive LPM specification (column 7 of Table 2).

Collectively, results included in Table 3 allow us building a sequence of potentially relevant relationships between our three parameters of interest. That is, a father's crime during high school raises the likelihood of dropping out, which in turn increases the probability of committing a crime when adult. In addition, the intergenerational transmission of crime itself is higher if pupils drop out from school. Interestingly, this may suggest that paternal crime could have both a direct effect on adult crime, via the standard intergenerational effect, and an indirect effect

[7] Following the first stage of the IPWRA estimation, we performed an overidentification test for covariate balance. Such test does not reject the null hypothesis that the weights produced by the first stage estimation balance the covariates, presenting a p-value of 0.293.

through high school dropout. Still, dropping out from school appears to play a comparatively larger role in affecting adult crime.

5.1 Estimates of the mechanisms

In the previous section, we have established the relevance of dropping out from school and having a convicted father in influencing adult crime, together with the existence of direct and indirect effects of paternal crime on adult crime. Here, we explore empirically three further mechanisms through which high school dropout might influence adult crime. These concern: (i) an early exposure to a criminal environment; the roles of (ii) cognitive skills and (iii) biological father's nurture, respectively.

Tables 4 and 5 examine the contribution of an early exposure to a criminal environment, using Add Health wave I contextual data on the total number of crimes committed in the county where individuals resided when attending school. Table 4 corresponds to the LPM estimated in Table 2 this time augmented by a series of dummy variables defining quartiles of county-level crime rates and their interactions with high school dropout. Column 1 of Table 4 isolates the effect of a father's conviction during high school on adult crime in the absence of any other control or fixed effects, while column 2 adds quartiles of crime rates together with school and state fixed effects. While the intergenerational effect of crime appears to be line with the previous LPM estimates in the absence of covariates, we now also observe a gradient in the effects of the third and fourths quartiles of the crime rate dummies. This may imply that individuals who were exposed to higher crime rates in their county (when at school) present an increased probability of committing a crime later on in life (9.5 and 10.8pp). Yet, the estimated coefficient of the variable defining the fourth quartile is only weakly statistically significant. Further models also account incrementally for individual-level variables, including high school dropout, and interactions between dropping out and quartiles of crime rates (column 3), as well as the full set of family-level observables (column 4). Although a gradient in the third and fourth quartiles of crime rates is still present, the estimated corresponding effects are now both weakly statistically significant. As expected, the interactions between the upper quartiles of crime rates and dropout are negative as they need to be interpreted together with the estimated baseline effect of school dropout. That is, living in counties with a higher concentration of crime (third and fourth quartiles of crime rates) appear to decrease the estimated probability of dropping out if compared to the baseline case of counties with the lowest level of crime. This could suggest that the higher the concentration of crime during high school, the lower the role of dropping out from school in influencing adult crime. Still, none of the estimated effects of such interactions are found to be strongly statistically significant. Overall, the results produced by our LPM models do not provide strong evidence in support of a long-term direct effect of a criminal environment on adult crime. Yet, we should keep in mind that our variables measure crime at county level and so they might not be able to identify effects within smaller geographical units such as neighbourhoods.

Corresponding IPWRA results in Table 5 show a statistically significant gradient in column 1, when we look at the determinants of dropout (treatment model). That is, living in counties with an increasing concentration of crime during high school, greatly enhances the likelihood of leaving school earlier (third and fourth quartiles, 32.2 and 41.9pp respectively). Estimates reported in columns 2 and 3 (outcome models), confirm that the intergenerational effect of crime is stronger for dropouts but do not show a perceivable gradient in the quartiles of crime rates, for either dropouts and non dropouts. The ATE of high school dropout on adult crime is still highly statistical significance and similar in size to that of our baseline IPWRA estimates (17.2pp). In sum, the empirical tests of this first mechanism show that while we find a short-term effect of an early exposure to a criminal environment on the probability of dropping out from high school, there is not enough evidence to conclude that this has also a residual long-term direct effect on adult crime.

Tables 6 and 7 analyse the contribution of cognitive skills. More specifically, we repeat the estimation of LPM and IPWRA models by excluding individuals who earned a General Education Degree (GED), after dropping out from school. We do so as GED holders have been found to present systematically higher cognitive skills if compared to pure dropouts (i.e. dropouts who did not subsequently earn a high school equivalent degree).[8] Results from the most comprehensive linear probability model reported in column 7 of Table 6 confirm that dropping out from school has an important effect on adult crime among pure dropouts (16.9pp). In addition, the estimated intergenerational correlation is also statistically significant (11.8pp, i.e. around 1pp larger if compared to the one produced by our baseline LPM). The latter result points towards a potentially stronger association of criminal behaviour across generations among individuals with lower levels of cognitive skills. The IPWRA model estimated on pure dropouts (excluding GED holders, Table 7) confirms an increased effect of a father's crime on school dropout itself (column 1) as well as statistically significant and larger intergenerational effects for dropouts (column 3). The ATE of dropping out from school on adult crime is also larger in size (21.8pp).

These estimates indicate that, among pupils with weaker cognitive skills, a father's conviction during high school is an even stronger determinant of school drop out. Moreover, early school attrition would put such dropouts at a higher risk of committing a crime in adulthood as well. Ultimately, this could reveal the relevance of cognitive skills in influencing both the relationship between human capital and crime and the intergenerational transmission of crime.

[8]As Heckman and Rubinstein (2001) suggest, GED holders are equated to standard high school graduates who do not pursue a college education, at least in terms of cognitive skills, as defined by an eight-hour subject-based test. Yet, both GED holders and pure dropouts appear to have similar lower levels of noncognitive skills.

Finally, Table 8 reports results focused on a third mechanism based on biological parents' nurture. In this case, we repeat the estimation of our LPM by excluding the presence of a biological father living with the pupil's family during high school (upper part of Table 8). This should enable exploring whether our key results about high school dropout and the intergenerational effect of crime would change in the absence of a (potentially convicted) biological father's nurturing influence.[9] Interestingly, whereas the effect of early school attrition is similar to the one produced by the baseline estimates in Table 2, the intergenerational correlation of crime of the most comprehensive specification reported in column 7 is still highly statistically significant but almost half the size (6pp vs 11pp). This might imply that the nurturing effect of a biological father accounts for a large fraction of the estimated correlation of crime across generations. In order to examine the broader influence of nurture on children's adult crime, we also test our results by imposing the absence of a biological mother (lower part of Table 8). In this case, both the effect of dropping out from school and the one of a father's conviction on adult crime are in line with previous estimates (column 7, 18.7pp and 10.4pp, correspondingly). This could suggest a weaker influence of a biological mother's nurturing effect on school dropout and the intergenerational transmission of crime.

5.2 Robustness Checks

In this section we consider a series of robustness checks to explore the validity of our main findings. Tables 9 and 10 repeat the estimation of our baseline LPM and IPWRA models using a stricter definition of adult crime. More specifically, adult crime is defined using a binary variable taking value 1 when pupils were convicted and spent time in jail at least once after 18 years old, 0 otherwise. Both sets of results reveal that the use of such dependent variable would not substantially change our results. Looking at LPM (Table 9), there is a still statistically significant and slightly larger effect of high school dropout on adult crime (compared to corresponding results in Table 2). In addition, the estimated intergenerational effect of crime is virtually identical in terms of statistical significance and magnitude (see column 7). Results of the corresponding IPWRA model reported in Table 10 are essentially confirmed as well, with larger effects of a father's conviction on school dropout and on adult crime among dropouts (columns 1 and 2). The ATE is also in line with previous results in Table 3.

Since our results appear to be very similar across alternative definitions of adult crime, we prefer relying on our initial definition leading to a larger sample and thus allowing the explaration of all potential mechanisms. However, and based on these results, we are aware that the estimates presented earlier may represent the lower bound of the corresponding effects of interest.

[9]Due to the reduced sample size, the corresponding IPWRA models could not be estimated.

As a final robustness check, we account for the role of noncognitive skills. We do this by reestimating our baseline models in Table 2 and 3, this time by excluding the three personality traits (i.e. conscientiousness, neuroticism, extroversion) otherwise included among the individual characteristics in previous models. Results, available upon request, are virtually identical, suggesting that personality traits might not be playing a fundamental role in our models.

6 Conclusions

We examine the relationship between high school dropout and pupils' adult crime by accounting for the role of the intergenerational transmission of crime. Given policy relevance and associated costs of early school attrition and crime, understanding the different pathways linking them is crucial for devising more targeted policies to curb both high school dropout and criminal behaviour. Our study is motivated by the lack of evidence on the potential direct and indirect effects of the intergenerational transmission of crime within the association between schooling and adult crime. Ultimately, our approach bridges the literature on the impact of human capital on crime with that concerned with the intergenerational transmission of crime.

We propose a simple human capital model describing pupils' choices around high school completion and adult crime, allowing these to be affected by paternal crime. This model also affords incorporating the influence of early exposure to high levels of crime; cognitive skills; and nurture. We test empirically the relationships highlighted by our theoretical model using a series of fixed effects linear probability models and inverse probability weights regression adjustment specifications. Importantly, the two-stage estimation of the latter enables identifying direct and indirect effects, via high school dropout, of the intergenerational transmission of crime. All models are estimated using rich information on education, crime and a host of socioeconomic and demographic variables drawn from Add Health.

Our empirical analysis allows shedding light on how education and fathers' criminal behaviour interact in affecting pupils' crime later on in life. Consistent with the literature, we find that the likelihood of committing a crime when adult is higher when a pupil drops out from school and when the pupil's father was convicted. Our specific contribution consists in showing the existence of an indirect mechanism whereby a father's conviction increases the probability of dropping out from school. The indirect mechanism of paternal crime operates in addition to its direct effect on a pupil's adult crime, i.e. the direct intergenerational effect, and it reinforces the likelihood of committing crime in adulthood. These findings thus suggest that a father's criminal behaviour has both a direct long-term effect on pupils' adult crime as well an indirect effect mediated by high school dropout. Accordingly, pupils with a family history of criminal behaviour may suffer from a double penalty: they are at higher risk of committing crimes in adulthood *per se* but they are also at higher risk of dropping out, which is an additional predictor of future crimes.

Among the broader determinants of adult crime, we find that whereas the exposure to high levels of crime during high school increases school dropout, it does does not appear to have a residual direct effect on adult crime. In addition, lower levels of cognitive abilities reinforce dropout and strengthen intergenerational effects. Finally, we provide suggestive evidence that the absence of a (potentially convicted) biological father reduces the persistence of crime across generations.

From a policy perspective, our results contribute to the debate centred around the identification of the mechanisms through which high school drop out affects crime later on in life and may have relevant implications. In particular, they allow identifying a specific category of individuals, i.e. pupils with convicted fathers, for whom programs aimed at preventing school attrition can be doubly beneficial. That is, not only such programs could improve pupils' future wellbeing and human capital but they may also decrease the likelihood that pupils will commit crimes later on in life.

Acknowledgments

Eugenio Zucchelli gratefully acknowledges financial support from the Tomás y Valiente Fellowship funded by the Madrid Institute for Advanced Study (MIAS), Universidad Autónoma de Madrid (UAM), Spain, as well as from grants PID2019-105688RB-I00 (Spanish Ministry of Science, Innovation and Universities), SI1/PJI/2019-00326 and H2019/HUM-5793 (Autonomous Community of Madrid). We would like to thank attendants to seminars at Universidad Autónoma de Madrid (UAM) and the Hertie School of Governance (Berlin) as well as participants to the 2020 Workshop in Applied Risk and Health Economics (Ca' Foscari, Venice) and the 2019 International Workshop on Pressing Issues in the Economics of Education (UAM, Madrid) for useful suggestions.

Bibliography

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. In *The Economic Dimensions of Crime*, pp. 13–68. Springer.

Becker, G. S. (1994). Human capital revisited. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, pp. 15–28. The University of Chicago Press.

Black, S. E., P. J. Devereux, et al. (2011). Recent developments in intergenerational mobility. *Handbook of Labor Economics 4*, 1487–1541.

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

Cattaneo, M. D. (2010). Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics 155*(2), 138–154.

Chalfin, A. (2015). Economic costs of crime. In *The Encyclopedia of Crime and Punishment*, pp. 1–12. Wiley Online Library.

Chetty, R. and N. Hendren (2018). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics 133*(3), 1107–1162.

Cook, P. J. and S. Kang (2016). Birthdays, schooling, and crime: Regression-discontinuity analysis of school performance, delinquency, dropout, and crime initiation. *American Economic Journal: Applied Economics 8*(1), 33–57.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. In *Handbook of the Economics of Education*, Volume 1, pp. 697–812. Elsevier.

Damm, A. P. and C. Dustmann (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review 104*(6), 1806–32.

Duncan, G., A. Kalil, S. E. Mayer, R. Tepper, and M. R. Payne (2005). The apple does not fall far from the tree. In *Unequal Chances: Family Background and Economic Success*, pp. 23–79. Princeton University Press, Princeton, NJ.

Dunn, L. and D. Dunn (2007). *Peabody Picture Vocabulary test, 4th edition.* Bloomington: NCS Pearson.

Fagan, J. and E. Pabon (1990). Contributions of delinquency and substance use to school dropout among inner-city youths. *Youth and Society 21*(3), 306–354.

Freeman, R. B. (1999). The economics of crime. In *Handbook of Labor Economics*, Volume 3, pp. 3529–3571. Elsevier.

Grogger, J. (1997). Local violence and educational attainment. *Journal of Human Resources 32*(4), 659–682.

Harris, K. M., F. Florey, J. Tabor, P. S. Bearman, J. Jones, and J. R. Udry (2009). *The National Longitudinal Study of Adolescent Health: Research Design.* Carolina Population Center, University of North Carolina, Chapel Hill, NC.

Heckman, J. J. and P. A. LaFontaine (2010). The american high school graduation rate: Trends and levels. *Review of Economics and Statistics 92*(2), 244–262.

Heckman, J. J. and Y. Rubinstein (2001, May). The importance of noncognitive skills: Lessons from the ged testing program. *American Economic Review 91*(2), 145–149.

Hjalmarsson, R., H. Holmlund, and M. J. Lindquist (2015). The effect of education on criminal convictions and incarceration: Causal evidence from micro-data. *Economic Journal 125*(587), 1290–1326.

Hjalmarsson, R. and M. J. Lindquist (2012). Like godfather, like son exploring the intergenerational nature of crime. *Journal of Human Resources 47*(2), 550–582.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

Iyer, S. (2016). The new economics of religion. *Journal of Economic Literature 54*(2), 395–441.

Lochner, L. (2004). Education, work, and crime: A human capital approach. *International Economic Review 45*(3), 811–843.

Lochner, L. and E. Moretti (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review 94*(1), 155–189.

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

McFarland, J., J. Cui, and P. Stark (2018). *Trends in High School Dropout and Completion Rates in the United States: 2014 (NCES 2018-117).* U.S. Department of Education.

Migali, G. and E. Zucchelli (2017). Personality traits, forgone health care and high school dropout: Evidence from us adolescents. *Journal of Economic Psychology 62*, 98–119.

Monrad, M. (2007). *High School Dropout: A Quick Stats Fact Sheet.* ERIC.

OECD (2019). *Education at a Glance 2019.*

O'Riordan, C. and M. O'Connell (2014). Predicting adult involvement in crime: Personality measures are significant, socio-economic measures are not. *Personality and Individual Differences 68*, 98 – 101.

Tauchen, H., A. D. Witte, and H. Griesinger (1994). Criminal deterrence: Revisiting the issue with a birth cohort. *Review of Economics and Statistics 76*(3), 399–412.

Thornberry, T. P., M. Moore, and R. Christenson (1985). The effect of dropping out of high school on subsequent criminal behavior. *Criminology 23*(1), 3–18.

Young, J. K. and A. A. Beaujean (2011). Measuring personality in wave I of the national longitudinal study of adolescent health. *Frontiers in Psychology 2*, 158.

Table 1: Summary statistics

Variable	Obs	Mean	Std. Dev.
Individual-level variables			
Adult crime	11634	0.227	0.419
Ever sentenced	10598	0.156	0.363
High school dropout	11634	0.078	0.268
Pure dropout (excluding GED holders)	11303	0.051	0.220
Father crime	11634	0.093	0.290
Male	11634	0.449	0.497
White	11634	0.531	0.499
Hispanic	11634	0.143	0.350
African-American	11634	0.219	0.413
Asian	11634	0.079	0.270
General ill-health	11634	0.066	0.248
Depression	11634	0.098	0.297
Peabody vocabulary test	11634	101.204	14.519
Learning disabilities	11634	0.095	0.293
Risky attitude	11634	0.335	0.472
Myopic behaviour	11634	0.125	0.331
Religious	11634	0.872	0.334
Risky health-behaviours	11634	0.226	0.418
Low conscientiousness	11634	0.144	0.352
Neuroticism	11634	0.077	0.267
Low extraversion	11634	0.118	0.322
Risky family health-behaviours	11634	0.384	0.486
Family-level variables			
Mother - no education	11634	0.036	0.186
Mother - middle primary education	11634	0.171	0.376
Mother - secondary education	11634	0.467	0.499
Mother - higher education	11634	0.279	0.448
Mother - unemployed/home	11634	0.190	0.392
Mother - routine/semiroutine job	11634	0.102	0.302
Mother - small employee/intermediate	11634	0.166	0.372
Mother - manager low/high	11634	0.214	0.410
Father - no education	11634	0.040	0.196
Father - middle primary education	11634	0.13	0.337
Father - secondary education	11634	0.349	0.477
Father - higher education	11634	0.252	0.434
Father - unemployed/home	11634	0.099	0.299
Father - routine/semiroutine job	11634	0.239	0.426
Father - small employee/intermediate	11634	0.068	0.252
Father - manager low/high	11634	0.151	0.358
Biological father absence	11627	0.204	0.403
Biological mother absence	11632	0.077	0.267

Table 2: Linear probability models: baseline estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	Adult crime						
High School dropout	0.238***	0.228***	0.181***				0.164***
	(0.014)	(0.015)	(0.015)				(0.015)
Father crime				0.153***	0.144***	0.115***	0.108***
				(0.013)	(0.014)	(0.013)	(0.013)
School and State FE		✓	✓		✓	✓	✓
Individual characteristics			✓			✓	✓
Family characteristics							✓
N	11634	11634	11634	11634	11634	11634	11634

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table reports baseline estimates obtained using linear probability models (LPM) on the effect of high school dropout on pupils' adult crime (columns 1-3) and the effect of fathers' convictions on pupils' adult crime (columns 4-6). All models include incrementally pseudo-state and school fixed effects and the full battery individual-level and family-level observed characteristics. Column 7 reports estimates of the most comprehensive linear probability model including the effects of both high school dropout and fathers' convictions.

Table 3: Inverse probability weighting regression adjustment: baseline estimates

	(1)	(2)	(3)
Dependent variable	High school dropout	Adult crime	
Sample	Full sample	Non Dropouts	Dropouts
Father crime	0.644***	0.116***	0.137***
	(0.105)	(0.051)	(0.051)
School and State FE	✓	✓	✓
Individual characteristics	✓	✓	✓
Family characteristics	✓	✓	✓
N	11634	10725	909
ATE		0.189***	
		(0.027)	

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table reports results of the two-stage estimation of the inverse probability weighting regression adjustment (IPWRA) baseline specification. Column 1 reports results of the first-stage probit model on the effect of fathers' convictions on high school dropout (treatment model). Columns 2 and 3 report results of the second-stage linear weighted models on the effects of fathers' convictions on pupils' adult crime for the sub-samples of non-dropouts and dropouts (outcome models). All models include incrementally pseudo-state and school fixed effects and the full battery individual-level and family-level observed characteristics. The average treatment effect (ATE) of high school dropout on pupils' adult crime is obtained as the difference between the predicted outcomes of the second-stage linear weighted regressions.

Table 4: Linear probability models: early exposure to crime

	(1)	(2)	(3)	(4)
Dependent variable		Adult crime		
Father crime	0.155***	0.145***	0.115***	0.109***
	(0.013)	(0.014)	(0.013)	(0.013)
County crime level (Q2)		-0.007	-0.005	-0.004
		(0.039)	(0.038)	(0.038)
County crime level (Q3)		0.095**	0.079*	0.079*
		(0.048)	(0.046)	(0.046)
County crime level (Q4)		0.108*	0.100*	0.095*
		(0.057)	(0.055)	(0.055)
High school dropout			0.193***	0.190***
			(0.029)	(0.029)
Dropout×County crime level (Q2)			0.040	0.036
			(0.043)	(0.043)
Dropout×County crime level (Q3)			-0.065	-0.066*
			(0.040)	(0.040)
Dropout×County crime level (Q4)			-0.058	-0.059
			(0.039)	(0.039)
School and State FE		✓	✓	✓
Individual characteristics			✓	✓
Family characteristics				✓
N	11346	11346	11346	11346

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table reports estimates obtained using linear probability models (LPM) accounting for the effect of the exposure to different levels of crime during high school on pupils' adult crime. Column 1 includes an estimate of the effect of fathers' convictions on adult crime; column 2 adds quartiles of county-level crime rates together with school and pseudo-state effects; column 3 also accounts for individual-level variables, including high school dropout, and interactions between high school dropout and quartiles of crime rates; column 4 adds the full set of family-level variables.

Table 5: Inverse probability weighting regression adjustment: early exposure to crime

	(1)	(2)	(3)
Dependent variable	High school dropout	Adult crime	
Sample	Full sample	Non Dropouts	Dropouts
Father crime	0.639***	0.117***	0.146***
	(0.105)	(0.015)	(0.052)
County crime level (Q2)	0.075	0.005	-0.094
	(0.163)	(0.015)	(0.059)
County crime level (Q3)	0.322**	-0.001	-0.126*
	(0.162)	(0.017)	(0.065)
County crime level (Q4)	0.419***	0.025	-0.094
	(0.157)	(0.015)	(0.065)
School and State FE	✓	✓	✓
Individual characteristics	✓	✓	✓
Family characteristics	✓	✓	✓
N	11346	10456	890
ATE		0.172***	
		(0.025)	

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table reports results of the two-stage estimation of the inverse probability weighting regression adjustment (IPWRA) specification accounting for the effect of the exposure to different levels of crime during high school on pupils' adult crime. Column 1 reports results of the first-stage probit model on the effect of fathers' convictions on high school dropout (treatment model). Columns 2 and 3 report results of the second-stage linear weighted models on the effects of fathers' convictions on pupils' adult crime for the sub-samples of non dropouts and dropouts (outcome models). All models include quartiles of county-level crime rates. All models include incrementally pseudo-state and school fixed effects and the full battery individual-level and family-level observed characteristics. The average treatment effect (ATE) of high school dropout on pupils' adult crime is obtained as the difference between the predicted outcomes of the second-stage linear weighted regressions.

Table 6: Linear probability models: pure dropouts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable				Adult crime			
High School dropout	0.246***	0.236***	0.185***				0.169***
	(0.018)	(0.018)	(0.018)				(0.018)
Father crime				0.157***	0.149***	0.134***	0.118***
				(0.014)	(0.014)	(0.013)	(0.013)
School and State FE		✓	✓		✓	✓	✓
Individual characteristics			✓			✓	✓
Family characteristics							✓
N	11303	11303	11303	11303	11303	11303	11303

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table reports estimates obtained using linear probability models (LPM) accounting for the effect of cognitive skills. Columns 1-3 report estimates of the effect of high school dropout on pupils' adult crime; columns 4-6 report estimates of the effect of fathers' convictions on pupils' adult crime; column 7 reports estimates of the most comprehensive linear probability model including the effects of both high school dropout and fathers' conviction on pupil's adult crime. All models include incrementally pseudo-state and school fixed effects and the full battery individual-level and family-level observed characteristics. All models were estimated on a sample of pure dropouts obtained by excluding General Educational Development (GED) certificate holders to account for the effect of cognitive skills.

Table 7: Inverse probability weighting regression adjustment: pure dropouts

	(1)	(2)	(3)
Dependent variable	High school dropout	Adult crime	
Sample	Full sample	Non Dropouts	Dropouts
Father crime	0.693***	0.115***	0.235***
	(0.134)	(0.014)	(0.048)
School and State FE	✓	✓	✓
Individual characteristics	✓	✓	✓
Family characteristics	✓	✓	✓
N	11303	10725	578
ATE		0.218***	
		(0.038)	

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table includes results of the two-stage estimation of the inverse probability weighting regression adjustment (IPWRA) specification accounting for the effect of cognitive skills. Column 1 reports results of the first-stage probit model on the effect of fathers' convictions on high school dropout (treatment model). Columns 2 and 3 report results of the second-stage linear weighted models on the effects of fathers' convictions on pupils' adult crime for the sub-samples of non dropouts and dropouts (outcome models). All models include incrementally pseudo-state and school fixed effects and the full battery individual-level and family-level observed characteristics. The average treatment effect (ATE) of high school dropout on pupils' adult crime is obtained as the difference between the predicted outcomes of the second-stage linear weighted regressions. All models were estimated on a sample of pure dropouts obtained by excluding General Educational Development (GED) certificate holders to account for the effect of cognitive skills.

Table 8: Linear probability models: nurture

Biological father absence	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable				Adult crime			
High School dropout	0.217***	0.204***	0.186***				0.179***
	(0.026)	(0.028)	(0.028)				(0.028)
Father crime				0.094***	0.085***	0.065***	0.060***
				(0.022)	(0.023)	(0.022)	(0.022)
School and State FE		✓	✓		✓	✓	✓
Individual characteristics			✓			✓	✓
Family characteristics							✓
N	2371	2371	2371	2371	2371	2371	2371

Biological mother absence	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Dependent variable				Adult crime			
High School dropout	0.231***	0.220***	0.218***				0.187***
	(0.044)	(0.049)	(0.050)				(0.051)
Father crime				0.150***	0.144***	0.113***	0.104**
				(0.039)	(0.042)	(0.041)	(0.042)
School and State FE		✓	✓		✓	✓	✓
Individual characteristics			✓			✓	✓
Family characteristics							✓
N	888	888	888	888	888	888	888

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table reports estimates obtained using linear probability models (LPM) accounting for the effect of parental nurture on pupils' adult crime by alternatively excluding the presence of a biological father and a biological mother living with the pupil's family during high school (upper and lower part of the table, respectively). Columns 1-3 and 8-10 report estimates on the effect of high school dropout on pupils' adult crime while columns 4-6 and 11-13 report results on the effect of fathers' convictions on pupils' adult crime. All models include incrementally pseudo-state and school fixed effects and the full battery individual-level and family-level observed characteristics. Columns 7 and 14 report estimates of the most comprehensive LPM including the effects of both high school dropout and fathers' convictions.

Table 9: Linear probability models: alternative definition of adult crime

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable				Adult crime			
High School dropout	0.238***	0.232***	0.191***				0.172***
	(0.013)	(0.014)	(0.014)				(0.014)
Father crime				0.156***	0.147***	0.130***	0.109***
				(0.012)	(0.012)	(0.012)	(0.012)
School and State FE		✓	✓		✓	✓	✓
Individual characteristics			✓			✓	✓
Family characteristics							✓
N	10598	10598	10598	10598	10598	10598	10598

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table reports estimates obtained using linear probability models (LPM) employing a stricter definition of pupils' adult crime (whether they were convicted and spent time in jail at least once after 18 years old). Columns 1-3 report results on the effect of high school dropout on pupils' adult crime and columns 4-6 report results on the effect of fathers' convictions on pupils' adult crime. All models include incrementally pseudo-state and school fixed effects and the full battery individual-level and family-level observed characteristics. Column 7 reports estimates of the most comprehensive linear probability model including the effects of both high school dropout and fathers' convictions.

Table 10: Inverse probability weighting regression adjustment: alternative definition of adult crime

	(1)	(2)	(3)
Dependent variable	High school dropout	Adult crime	
Sample	Full sample	Non Dropouts	Dropouts
Father crime	0.680***	0.116***	0.162***
	(0.103)	(0.015)	(0.054)
School and State FE	✓	✓	✓
Individual characteristics	✓	✓	✓
Family characteristics	✓	✓	✓
N	10598	9809	789
ATE		0.196***	
		(0.026)	

Notes: Source: Add Health. Significance levels: *** 1%, ** 5%, *10%; standard errors in parentheses. This table includes results of the two-stage estimation of the inverse probability weighting regression adjustment (IPWRA) specification employing a stricter definition of adult crime (whether were they convicted and spent time in jail at least once after 18 years old). Column 1 reports results of the first-stage probit model on the effect of fathers' convictions on high school dropout (treatment model). Columns 2 and 3 report results of the second-stage linear weighted models on the effects of fathers' convictions on pupils' adult crime for the sub-samples of non-dropouts and dropouts (outcome models). All models include incrementally pseudo-state and school fixed effects and the full battery individual-level and family-level observed characteristics. The average treatment effect (ATE) of high school dropout on pupils' adult crime is obtained as the difference between the predicted outcomes of the second-stage linear weighted regressions.

Appendix

Solving the model

Recall that c and l refer to the time allocated to criminal activities and legitimate jobs in period 2, respectively, and that s and n is the time allocated to schooling and non schooling activities in period 1; H denotes human capital, and X describes a vector of exogenous variables. We solve the model by backward induction.

In period 2, the optimal allocation of time between criminal and legitimate activities solves

$$\max_{c,l} U_2\left(y_2; X\right) \tag{12}$$

$$\text{s.t. } y_2 = m_c\left(c; H, X\right) + m_l\left(l; H, X\right) \tag{13}$$

$$T_2 = c + l \tag{14}$$

$$H = h\left(s; X\right) \tag{15}$$

Hence optimal time devoted to crime satisfies (omitting the arguments)

$$\Delta \equiv \frac{\partial m_c}{\partial c} - \frac{\partial m_l}{\partial l} \geq 0, \tag{16}$$

where Δ is the gap between the marginal income of crime and the marginal income of labour. For analytical simplicity, we will assume that the above condition holds with equality. Hence, optimal time used for criminal activities $c^* = c\left(H, X\right)$ makes the productivity gap Δ between criminal and legitimate activities equal to zero.

In period 1, the optimal allocation of schooling and non schooling time solves

$$\max_{s,n} U_1\left(y_1; X\right) + \beta U_2\left(y_2; X\right) \tag{17}$$

$$\text{s.t. } y_1 = m\left(n; X\right) \tag{18}$$

$$y_2 = m_c\left(c^*; H, X\right) + m_l\left(T_2 - c^*; H, X\right) \tag{19}$$

$$T_1 = s + n \tag{20}$$

$$H = h\left(s, X\right) \tag{21}$$

Assuming an internal solution exists, the associated first order condition is

$$\frac{\partial U_1}{\partial y_1}\frac{\partial m}{\partial n} = \beta \frac{\partial U_2}{\partial y_2}\frac{\partial y_2}{\partial H}\frac{\partial H}{\partial s} \tag{22}$$

Hence, optimal schooling time s^* depends on the trade off between the utility that can be obtained by not investing time on schooling, and the impact of higher human capital on future income (either from labour or crime) and, ultimately, future discounted utility.

The impact of human capital on the optimal allocation of time between crime and a legitimate job is

$$\frac{dc^*}{dH} = \xi \frac{\partial \Delta}{\partial H} \gtreqless 0 \tag{23}$$

where $\frac{\partial \Delta}{\partial H} \equiv \frac{\partial}{\partial H}\left(\frac{\partial m_c}{\partial c} - \frac{\partial m_l}{\partial l}\right)$ describes the effect of human capital on the productivity gap between criminal and non criminal activities, and $\xi = -\left(\frac{\partial^2 m_c}{\partial c^2} + \frac{\partial^2 m_l}{\partial l^2}\right)^{-1} > 0$ by concavity. This expression means that education will decrease the time allocated to criminal activities if and only if human capital increases the marginal productivity of a legal job more than the marginal productivity of a criminal job.

We can assess the impact of the exogenous variables described by $X^k \in X$, $k \in \{1, 2, ..., K\}$ on crime in period 2 by computing

$$\begin{aligned} \frac{dc^*}{dX^k} &= \xi \left(\frac{\partial \Delta}{\partial H}\frac{\partial H}{\partial X^k} + \frac{\partial \Delta}{\partial X^k}\right) \\ &= \frac{dc^*}{dH}\frac{\partial H}{\partial X^k} + \xi \frac{\partial \Delta}{\partial X^k} \gtreqless 0 \end{aligned} \tag{24}$$

As shown in equation 23, the sign of $\frac{\partial \Delta}{\partial H}$ determines whether more schooling leads to less crime. The impact of the term $\frac{\partial H}{\partial X^k}$ describes how factor X^f directly affects the accumulation of human capital, while the third term $\frac{\partial \Delta}{\partial X^k}$ describes the impact of X^k on the crime-labour productivity gap.