Separating State Dependence, Experience, and Heterogeneity in a Model of Youth Crime and Education

Maria Antonella Mancino, Salvador Navarro, and David A. Rivers*

Department of Economics, University of Western Ontario, 1151 Richmond Street N., London, Ontario, N6A 5C2, Canada

June 30, 2016

Abstract

We study the determinants of youth crime using a dynamic discrete choice model of crime and education. We allow past education and criminal activities to affect current crime and educational decisions. We take advantage of a rich panel dataset on serious juvenile offenders, the Pathways to Desistance. Using a series of psychometric tests, we estimate a model of cognitive and social/emotional skills which feed into the crime and education model. This allows us to separately identify the roles of state dependence, returns to experience, and heterogeneity in driving crime and enrollment decisions among youth. We find small effects of experience and stronger evidence of state dependence and heterogeneity for crime and schooling. We provide evidence that, as a consequence, policies that affect individual heterogeneity (e.g., social/emotional skills), and those that temporarily keep youth away from crime, can have important and lasting effects even if criminal experience has already accumulated.

Keywords: crime, education, youth, dynamics

JEL codes: I21, I26, K42

*Corresponding author. Tel.: +1 519 661 2111 ext. 80446.
E-mail addresses: mmanci3@uwo.ca (M.A. Mancino), snavarr@uwo.ca (S. Navarro), drivers2@uwo.ca (D. A. Rivers)

1 Introduction

Empirical evidence suggests that youth account for a large share of crime. In the United States, 1.9 million youth between the ages of 15 and 19 were arrested in 2010, accounting for 19% of all arrests, despite representing only 7% of the total population.[1] Furthermore, numerous studies have found that criminal activity is highly persistent over time (Blumstein, Farrington, and Moitra, 1985; Nagin and Paternoster, 1991, 2000). This implies that reducing youth crime can have not only immediate effects on criminal activity, but also lasting effects as these individuals transition to adulthood.[2]

In order to design crime-reducing policies that effectively target youth, it is important to understand the determinants of youth crime. Recently there has been an increased recognition in the literature that education may be an important driver of criminal behavior. Increased educational attainment may increase future wages, which increases the return to legitimate work and raises the opportunity cost of illegal activities (Freeman, 1996; Lochner, 2004). Schooling may alter people's preferences, for example by increasing patience or risk aversion (Becker and Mulligan, 1997; Usher, 1997). By emphasizing social and emotional development, education can affect psychic or financial rewards from crime (Lochner, 2011a). Schooling can also have an incapacitation effect (Lochner, 2004; Jacob and Lefgren, 2003), or it can cause increased criminal activity by increasing the concentration of young people, leading to more violent confrontations (Jacob and Lefgren, 2003) or increased drug-related offenses by bringing together buyers and sellers.[3] Schooling can affect social networks, and these networks could influence criminal behavior, for example via gang participation (Lochner, 2010).

There are also channels through which crime can affect educational decisions. Having a criminal record may reduce the probability of obtaining a legitimate job, or may reduce the expected wage, lowering the returns to education (Hansen, 2011; Kim, 2014). Criminal experience may also increase the returns to criminal activity, thus lowering the relative returns to legitimate work and therefore education (Loughran et al., 2013; Munyo, 2015). This could, in turn, feed back into crime choices.

In this paper we study the determinants of youth crime in the context of a joint dynamic discrete choice model of crime and education, by allowing previous decisions to affect current choices. Understanding the relationships between crime and education has important policy implications. To

[1] These figures are based on data from the U.S. Census and the FBI's Uniform Crime Reports.

[2] In addition to the direct benefits to society of reducing crime, there are also indirect benefits. Research has found that incarceration negatively affects future earnings of individuals (Grogger, 1995, 1998; Waldfogel, 1994; Nagin and Waldfogel, 1995; Kling, 2006). Moreover, higher levels of crime have been found to reduce incentives for investment (Zelekha and Bar-Efrat, 2011).

[3] The literature is inconclusive on the direction of the effect of contemporaneous education on crime. Farrington et al. (1986), and Witte and Tauchen (1994) find that time spent at school is associated with lower levels of criminal behavior. Jacob and Lefgren (2003) and Luallen (2006) find that being in school causes a drop in property crime, but an increase in violent crime. Anderson (2014) finds that enrollment is negatively associated with both property and violent crime rates.

the extent that education and crime interact, this provides additional instruments for policy makers interested in reducing crime and/or increasing educational attainment.

The data we employ comes from the Pathways to Desistance (PD), a multi-site longitudinal study of serious adolescent offenders as they transition from adolescence into early adulthood. The Pathways to Desistance was designed specifically to study questions related to the evolution of criminal behavior, taking special care to also measure educational decisions and outcomes. As a result, the dataset contains a rich panel of information about decisions to participate in crime and enroll in school. This allows us to construct the criminal history of an individual as well as his educational experience and enrollment decisions over time. Each study participant was followed for a period of seven years after entering the survey which results in a comprehensive picture of life changes in a wide array of areas over the course of this time.[4] These features make the Pathways to Desistance data well-suited for understanding the dynamics in crime and education.

The relationship between crime and education has been studied using a variety of datasets, including the NLSY79 (Grogger, 1998; Lochner and Moretti, 2004; Lochner, 2004), the NLSY97 (Merlo and Wolpin, 2015), the Philadelphia Birth Cohort Study (Imai and Krishna, 2004; Tauchen, Witte, and Griesinger, 1994), the National Youth Survey (Imai, Katayama, and Krishna, 2006), and the National Longitudinal Study of Adolescent Health (Mocan and Rees, 2005), among others. A common feature of these datasets is that they study subsets of the population at large, and include very few serious offenders.

An advantage of studying only serious offenders through the PD data is that, to the extent that there is unobserved heterogeneity that leads some individuals to become serious offenders, we are more likely to be observing individuals who are on a criminal trajectory (Nagin and Land, 1993; Nagin, Farrington, and Moffitt, 1995). For policy makers interested in reducing overall crime rates, particularly violent crime rates, data on these serious offenders, who contribute significantly to aggregate crime rates, is necessary. While selecting on serious offenders has its advantages, one limitation is that we cannot necessarily generalize our findings to the population at large. The data are also less useful for studying the transition to becoming a serious offender, as we only observe those individuals that have already done so.

Our extremely rich set of control variables allows us to separate the effects of experience (captured by the accumulation of education and crime) from contemporaneous effects of education on crime, and from the effects of individual heterogeneity. Furthermore, we are able to separately account for the effects of state dependence in these decisions (captured by lagged decisions). Being able to separate

[4]We describe the dataset in more detail in Section 2.

these channels is important for evaluating potential policies aimed at either reducing crime or increasing educational attainment. For example, if there are large positive returns to criminal experience, then interventions to reduce crime need to be taken at early ages before experience accumulates. If instead the returns to experience are low, but there is a high degree of state dependence, then policies can be impactful at any age, but need to be repeated as the lagged effects depreciate.

The PD data includes a much larger set of targeted control variables than is typically available. In addition to standard socio-economic variables and information about individuals' families, the dataset also contains a number of additional individual-level variables that are particularly useful for our analysis. In each year the data contain a measure of each individual's perception about their probability of being caught if they commit a crime.[5] It also has information about drug usage, involvement in crime by family members, and a measure of how each individual discounts future events, among others.

An additional benefit of this dataset is that individuals are given a series of tests designed to measure unobserved heterogeneity, namely cognitive and social/emotional skills. Numerous studies have established that cognitive ability is a strong predictor of schooling attainment and wages (Cawley, Heckman, and Vytlacil, 2001; Murnane, Willett, and Levy, 1995), as well as a range of social behaviors (Herrnstein and Murray, 1994). Recently, an emerging body of research shows the effects of social/emotional skills (sometimes referred to as "non-cognitive ability") on outcomes such as labor market participation, health, and test scores (Heckman, Stixrud, and Urzua, 2006; Chiteji, 2010; Cobb-Clark and Tan, 2011). Focusing specifically on crime, Hill et al. (2011) show that programs targeting psychological factors besides cognitive ability were effective at reducing delinquency. Heckman, Stixrud, and Urzua (2006) show that both cognitive and non-cognitive skills influence a wide variety of risky activities such as smoking by age 18, imprisonment, and participation in illegal activities. Research from criminology and psychology has also found significant correlations between IQ, measures of personality, and crime/delinquency (Caspi et al., 1994; Agnew et al., 2002).

Incorporating these additional measures of observed and unobserved heterogeneity not only aids in separately identifying the various channels driving observed crime and education decisions. They also represent additional potential instruments for policy makers. To the extent that behavioral problems or drug use affect criminal activity, this provides additional opportunities to affect criminal behavior among youth by reducing drug use and/or improving social/emotional skills.

As a preview of our results, we find that measures of individual heterogeneity are important in explaining the patterns of enrollment and crime choices. In particular, many of the measures less

[5] Empirical estimates of crime deterrence based on the perceived certainty or severity of punishment on crime provide mixed results (Lochner, 2007; Paternoster and Simpson, 1996; Bachman, Paternoster, and Ward, 1992; Pogarsky and Piquero, 2003).

commonly observed in datasets, such as drug use, involvement in crime by family members, attitudes towards the future, and social/emotional skills, have some of the largest effects. We also find evidence of important dynamics. State dependence leads to the strongest effects, but there is evidence of small returns to experience.

The rest of the paper is organized as follows. In Section 2 we describe our data from the Pathways to Desistance. Section 3 contains our joint dynamic discrete choice model of crime and education. Section 4 presents the empirical results from our model, as well as a number of robustness checks. In Section 5, we provide some simulations of our model to illustrate how the enrollment and criminal behavior evolve over time and discuss some policy implications. Section 6 concludes.

2 Data

Our data come from the Pathways to Desistance (PD) study, a longitudinal investigation of the transition from adolescence to young adulthood for serious adolescent offenders.[6] Participants in the PD study are adolescents who were found guilty of a serious criminal offense (almost entirely felony offenses, but also including misdemeanor weapons offenses) in the juvenile or adult court systems in Maricopa County, Arizona, or Philadelphia County, Pennsylvania, between November 2000 and January 2003.[7] The study follows youth who were at least 14 years old and under 18 years old at the time of their offense. Individuals had to provide informed assent or consent to participate in the study.[8] Due to resource constraints and a cap of drug offenses, about one-half of those that met the age and offense requirements were approached to participate in the study.[9] In the end, 1,354 participants enrolled, yielding an enrollment rate of 67%.

The initial (baseline) survey occurred when individuals first entered the sample. For those in the juvenile system, the baseline interview was completed within 75 days after their adjudication, and for those in the adult system within 90 days after their decertification hearing (in Philadelphia) or arraignment (in Phoenix). There were six semi-annual follow-up interviews, followed by four annual follow-up interviews. They were typically conducted in the participant's home, or in a residential facility if the individual was in a jail or juvenile detention center. In total, the survey covers each individual for eight years. Individuals were paid $50 to participate in the baseline survey, with compensation increasing

[6]For more information on the Pathways to Desistance study see Schubert et al. (2004); Mulvey and Schubert (2012).

[7]We follow the terminology from the PD survey and interchangeably refer to Maricopa County as Phoenix, given that Phoenix is the main city within the county.

[8]Parental consent was obtained for all youth younger than 18 at the time of enrollment in the survey.

[9]The proportion of male youth found guilty of a drug charge was capped at 15% to avoid an overrepresentation of drug offenders. All female juveniles meeting the age and adjudicated crime requirements and all youths whose cases were being considered for trial in the adult system were eligible for enrollment, even if the charged crime was a drug offense.

for the follow-ups to minimize attrition (Monahan et al., 2009). The retention rate, measured as the share of participants completing a particular interview wave, was above 90% for the first six waves and no less than 83% for the following annual interviews.

One key feature of the PD data is that it follows individuals making school enrollment and crime decisions over time. This is a crucial feature for understanding the importance of dynamics in decisions about both crime and education. A second key feature of this dataset is that it contains extremely detailed data on individual characteristics that may be important for predicting both schooling and criminal activity.

The baseline survey contains basic demographic information including age, gender, ethnicity, and location (i.e., Maricopa or Philadelphia County). Additionally, the survey records the number of siblings, the number of children each individual has, whether individuals live with both natural parents[10], and whether any family members are involved in criminal activities.[11] We also observe whether individuals use drugs, as well as their perceived risk to offending (i.e., the individual-specific perceived probability of getting caught).[12] Furthermore, we have a measure of how much individuals care about the future, through a variable called the *Future Outlook Inventory*. This measure is created based on survey questions related to the assessment and implications of future outcomes and consideration of future consequences. Higher scores indicate a greater degree of future consideration and planning, and thus are associated with higher discount factors (lower discount rates).

Information on family criminal activities, number of children, the perceived risk to offending, drug use, and future outlook inventory is collected again in each follow-up survey. We supplement this information with data from the Bureau of Labor Statistics on local annual unemployment rates, data on the number of high schools from the National Center of Education Statistics, and data on the number of people between the ages of 15 and 19 in each county from the U.S. Census.[13]

In addition to the detailed information about observable characteristics of each individual, the PD data also contains the results from a large number of standard psychometric tests that were given to each person. These tests are designed to measure characteristics of the individual that we typically consider to be not directly observable, such as intellectual ability (e.g., IQ) and social/emotional capabilities

[10] Dornbusch et al. (1985) show that family composition during childhood may affect criminal behavior.

[11] Both criminal behavior and enrollment decisions of children can be affected by the criminal involvement of their parents as the social environment in the family becomes more unstable (Geller et al., 2009).

[12] The perceived risk is measured in each period by asking individuals how likely it is that they will be caught and arrested conditional on committing a particular crime. There are seven underlying measures, corresponding to each of the following crimes: fighting, robbery with a gun, stabbing someone, breaking into a store or home, stealing clothes from a store, vandalism, and auto theft. Response options ranged from 0 (no chance) to 10 (absolutely certain to be caught). Only the average across these seven responses is reported in the data.

[13] We use the latter two to compute the number of schools per person of high school age in each county-year pair, as a measure of the cost of attending school.

(e.g., impulse control, self-esteem, and ability to suppress one's aggression). We group these tests into those designed to measure cognitive skills and those designed to measure social/emotional skills. The cognitive tests are given only in the baseline survey, whereas the social/emotional tests are repeated in the follow-up surveys as well.

The cognitive measures include the *Wechsler Abbreviated Scale of Intelligence* (WASI) test score, which produces an estimate of general intellectual ability (IQ) based on two components: Vocabulary and Matrix Reasoning. In addition, we have two batteries of tests related to cognitive dysfunction: the *Stroop Color-Word Test* and the *Trail-Making Test.* The Stroop Color-Word Test is used to examine the effects of interference on reading ability, and the Trail-Making test is a measure of general brain function. The Stroop test has three parts, which relate to interference from colors, words, and both words and colors together. Subjects are asked to identify colors based on the written name of the color, or the color of the ink the word is printed in. The Trail-Making test measures general brain development and damage. It consists of two parts: Part A involves a series of numbers that the participant is required to connect in sequential order; Part B involves a series of numbers and letters and the participant is required to alternately connect letters and numbers in sequential order.

We also have several measures of social/emotional skills. First, the *Weinberger Adjustment Inventory* (WAI) is an assessment of an individual's social/emotional adjustment within the context of external constraints. The test is divided into three areas: impulse control, suppression of aggression, and consideration of others. Second, the *Psychosocial Maturity Inventory* (PSMI) provides measures of self-reliance, identity (i.e., self-esteem and consideration of life goals), and work orientation (i.e., pride in the successful completion of tasks).[14]

Finally, the dataset contains information on the enrollment and criminal activity decisions of each individual. In each survey, individuals are asked whether they have been enrolled in school during the recall period (either six-months or one year in length). In addition, in the baseline survey they are asked what is the highest grade that they have completed. We combine this variable with subsequent enrollment decisions to construct a measure of years of accumulated education in each year.

The data on criminal activity comes from self-reporting by each individual. The self-reported offenses (SRO) consist of 24 components, each of which relates to involvement in a different type of crime, e.g., destroying or damaging property, setting fires, or selling drugs. For each item, a set of follow-up questions collect more information regarding the reported offense (e.g., "how many times have you done this in the past N months?") and can be used to identify whether the adolescent reports committing an act within the recall period, the frequency of these acts, as well as whether the act was committed

[14]In both the WAI and PSMI tests, higher scores indicate more positive behavior.

alone or with a group. The baseline questionnaire also collects information on the subject's age at the first time he engaged in each criminal activity.

For our analysis we combine these crime components into three categories: (i) **violent crime**, which consists of those offenses involving force or threat of force (e.g., robbery and assault), (ii) **property crime**, which includes burglary, larceny-theft, motor vehicle theft, and arson; and (iii) **drug-related crime** (e.g., selling marijuana or other drugs). While violent crime typically also includes murder and rape, these crimes are not reported in our data due to confidentiality restrictions.[15] Our main results are based on one aggregate category, by combining all three sub-categories.

Although self-reported crime may suffer from under-reporting, it is the most direct measure of criminal participation available. It includes all crimes committed by the individual, and not just those for which the individual was caught. In order to encourage accurate self-reporting, individual responses are kept confidential, and participants were given a certificate of confidentiality from the U.S. Department of Justice. Furthermore, in our analysis we only use information on whether an individual has engaged in a criminal activity, and not the intensity. This does not require that people truthfully report the extent of their criminal activities, only that they accurately report criminal participation.

While we have data on the criminal activities of each individual once they enter the survey, we generally do not know their criminal history prior to the initial survey, with the exception of knowing the age at which each individual first committed each of the crimes.[16] In order to deal with this missing data problem, we impute the years of crime using the following procedure. We first estimate a probit model for crime using the data on age and the time-invariant covariates (ethnicity, location, gender, intact family, number of siblings) as regressors. This gives us an estimate of the probability of crime in each period, conditional on age and time-invariant characteristics. Combined with the age of first crime variable, we can then estimate the expected number of years of crime by the time the individual enters the baseline survey. Experience in subsequent years is then calculated based on this estimate and on the observed crime decisions.[17]

We construct four panel datasets, one for each of the three crime measures described above and

[15] Not all of the components are mapped into one of our three categories, e.g., example drunk driving and carrying a gun. In total we use 16 of the 24.

[16] For some individuals we can infer their entire criminal history, for example those whose first crime triggered their entry into the survey.

[17] An alternative to our imputation procedure is, at estimation time, to use the probabilities predicted by our model in Section 3 to integrate the likelihood for each individual against the distribution of unobserved criminal experience. As a robustness check, we estimated our model using this alternative approach to deal with the unobserved criminal experience. Specifically, for each individual in our dataset, we simulate S possible paths of crime and enrollment decisions from age at first crime to age of entry into the survey, by sampling S draws of the errors in the crime and enrollment equations. For each individual we then calculated S likelihoods, corresponding to each of the S simulated paths. The individual contribution to our overall likelihood is calculated as the average (i.e., the Monte Carlo integral) over these paths. The results were very similar to our baseline estimates, and since this procedure substantially increased the computational burden, we decided not to use it over our simpler imputation procedure.

one with all crime together. Each panel includes all individuals for whom all the relevant variables are reported. The panels are constructed using annual data. Individuals are included in the dataset until at least one of the relevant variables is missing for a given year (i.e., an unbalanced panel). Under this procedure, we are left with 1,168, 1,188, 1,191 and 1,187 individuals in the drug-crime, violent, property and overall crime panels, respectively.[18] Each sample includes, at most, eight years for each individual. The attrition rate in the overall crime sample is on average slightly less than 6% per year.[19]

Table 1 reports descriptive statistics for our four samples. There are several statistics that we wish to highlight. First, crime rates in the sample (i.e., the fraction of individual-year pairs in which a crime was committed) are quite high. The violent crime rate is 44%, 29% for property crime, and 21% for drug related crime. These high crime rates (particularly for violent crime) come from the fact that all individuals in the dataset have been convicted of a serious criminal offense at least once, as this is a requirement for entering the dataset. About 14% of the sample is female. There is a large percentage of minorities, with blacks and Hispanics representing 40% and 34% of the sample, respectively. Drug use is also quite prominent, with an average of 47%. The average age for the first crime is 10.7 for violent, 11.5 for property, and 13.9 for drug-dealing crimes, illustrating that many of these adolescents start participating in criminal activities well before high school, particularly for violent and property crime.

Table 2 reports descriptive statistics for the tests designed to measure cognitive skills. In our empirical analysis we use the two components of IQ separately: the raw WASI Vocabulary Score and the raw WASI Matrix Reasoning Score. However, for interpretability, we report information on the distribution of IQ scores here as well. On average, IQ scores in our sample are substantially below the average score in the general population (100). In fact, almost 90% of individuals have a score below 100. For our measures of cognitive impairment, the Trail-Making scores take one of four values, where the lowest two values indicate either mild/moderate impairment or moderate/severe impairment. In our sample, 21% have some level of cognitive impairment according to Trail-Making A, and 38% under Trail-Making B. The Stroop Test scores take a continuum of values. For each test, scores above 40 are considered "normal". For the Color, Word, and Color/Word tests respectively, 52%, 36%, and 21% had scores below normal.

[18]The sample size in the overall crime sample is not necessarily the largest across all four samples. For instance, an individual can have missing data for violent crime and specifically state no involvement in property and drug crime. In this case, he is included in the property and drug crime samples as someone who did not commit crime, but dropped from the violent crime sample. For the overall sample, we do not know whether he committed a crime or not (since violent crime is missing), so he is dropped from the overall sample as well. In cases in which the individual has missing data for a certain crime category but expresses criminal engagement in any other specific crime category, then he is included in the overall sample, since it is clear that he participated in at least one type of crime.

[19]Special efforts were made to reduce attrition. Unless the participants explicitly withdrew from the study or died, interviewers continued to attempt to contact a research participant for future interviews even after one or more of the previous time-point interviews was missed. In addition, participants were paid on a graduated payment scale designed to encourage continued participation.

The raw social/emotional test scores are harder to interpret. In both the WAI and PSMI, individuals are given a set of questions and asked to indicate the extent to which the statement is true or false (WAI) on a scale of 1-5, or to what extent they either agree or disagree with the statement (PSMI) on a scale of 1-4. In both tests, responses are coded such that higher numbers indicate more positive behavior. For the section of the WAI measuring impulse control, 40% of the scores are below 3, indicating undesirable behavior. For suppression of aggression and consideration of others, the corresponding percentages are 50% and 18%, With the PSMI, the percentage of scores consistent with undesirable behavior (scores below 2.5), were considerably smaller: 5% (self reliance), 4% (identity), 15% (work orientation).

Figures 1-3 illustrate some of the key relationships in the data that our model seeks to explain: in particular, the contemporaneous and dynamic correlations between the education and crime decisions. Since age is highly correlated with both enrollment and crime decisions, we illustrate all of these relationships conditioning on age.

Figure 1 shows how the probability of committing crime depends on the lagged crime decision, and how this evolves with age.[20] Figure 2 shows the same for education. There are two important relationships to notice in these figures. First, both crime and education decisions are highly persistent in that individuals who committed crime (enrolled in school) in the previous period are much more likely to commit crime (enroll in school) in the current period. Second, there is some evidence of dynamic selection since, as individuals age, this relationship becomes stronger.

Figures 1 and 2 demonstrate strong persistence in crime and education decisions. What cannot be determined from the figures alone is the cause of this persistence (Heckman, 1981). This could be generated by persistent differences across individuals that are correlated with education and crime decisions. For example, it may be that low-skill youth are less likely to enroll in school and more likely to commit crimes. A second explanation is that there is state dependence in these decisions. For example, attending school may be easier if the individual has learned the previous year's material. A third possibility is that there are returns to previous experience. It may be the case that individuals become better at committing crimes with more practice, which increases the future probabilities of committing crimes. In our empirical analysis we attempt to disentangle all three potential causes for the observed persistence in decisions.

Figure 3 illustrates the contemporaneous link between youth crime and enrollment, suggesting a negative correlation in the mid teenage years. While this would seem to suggest a negative effect of

[20] There is a small number of individuals with lagged crime equal to zero at age 15. Since individuals with non-missing values of lagged crime at age 15 entered the sample the previous year, and since committing a crime is what triggers entry into the sample, we would expect all of the people to have lagged crime equal to one. However, individuals can be considered for the study if they are found guilty of a misdemeanor weapons crime, which we do not categorize into one of our crime types (violent, property, drug).

enrollment on crime, these results do not control for any heterogeneity (except age) across individuals that could also be driving this relationship. In addition, negatively correlated shocks to the enrollment and crime decisions could also generate this relationship. In the next section we present our model, and show how we are able to separately identify these confounding effects in order to recover the causal effect of enrollment on crime.

3 Model

Consider the problem of individuals indexed by i who decide at each age t whether or not to enroll in school and/or commit crime. The education choice is coded as $e_{i,t} = 1$ if the person goes to school in that period and 0 otherwise, and similarly for the crime choice $c_{i,t}$. The net utility of getting education in period t is a function of all relevant decision variables including lagged crime and enrollment decisions, years of crime and years of education up to t ($yc_{i,t}$ and $ye_{i,t}$), and a set of individual-specific characteristics ($z^e_{i,t}$,$z^c_{i,t}$) corresponding to the enrollment and crime equations, respectively:

$$v^e_{i,t} = z^e_{i,t}\beta^e + e_{i,t-1}\kappa^e + yc_{i,t}\lambda^e + ye_{i,t}\alpha^e + \eta^e_{i,t}, \tag{1}$$

where $\eta^e_{i,t}$ denotes unobservable individual-specific utility terms. An individual chooses to enroll in school ($e_{i,t} = 1$) if and only if $v^e_{i,t} > 0$.

Similarly, the crime choice is denoted as $c_{i,t} = 1$ if a crime is committed and 0 otherwise. The net utility of crime commission given the enrollment decision, is

$$v^c_{i,t} = z^c_{i,t}\beta^c + e_{i,t}\gamma^c + c_{i,t-1}\pi^c + yc_{i,t}\lambda^c + ye_{i,t}\alpha^c + \eta^c_{i,t} \tag{2}$$

where $\eta^c_{i,t}$ denotes unobservable individual-specific utility terms. Given the enrollment decision, the individual chooses to commit crime ($c_{i,t} = 1$) if and only if $v^c_{i,t} > 0$.

Notice that in equations (1) and (2) above, we allow contemporaneous enrollment to affect the crime decision, but not the other way around. The reason for this is that if we were to allow for both types of feedback effects, the resulting model would not be identified due to the problem of incoherency.[21] Therefore, we impose what is referred to in the literature as the coherency condition, by restricting the contemporaneous effect of crime on education to be zero.[22]

[21] See Heckman (1978) and Lewbel (2007) for further discussion of the identification problems associated with dummy endogenous variables in simultaneous equations models.

[22] We focus on this case because the literature is focused more on the effect of education on crime, as opposed to the effect of crime on education. Alternatively we could assume that the contemporaneous effect of enrollment on crime is zero. In Appendix Table A4, we provide results from the model with the contemporaneous effect in the other direction

Imposing the coherency condition makes our model triangular, which allows us to factor the likelihood in the following way:[23]

$$\Pr(c_{i,t}, e_{i,t}) = \Pr(c_{i,t} \mid e_{i,t}) \Pr(e_{i,t}),$$

where $\Pr(c_{i,t} = 1 \mid e_{i,t}) = \Pr\left(v^c_{i,t} > 0 \mid e_{i,t}\right)$ and $\Pr(e_{i,t} = 1) = \Pr\left(v^e_{i,t} > 0\right)$, and similarly for the probabilities of $c_{i,t} = 0$ and $e_{i,t} = 0$. If we were to assume that the errors in equations (1) and (2) are independent and normally distributed, we could estimate the model parameters by estimating separate probits. However, the assumption that the residuals are independent is unlikely to be true, as many of the factors driving enrollment decisions are likely to drive crime decisions as well. When this is the case, $e_{i,t}$ will be endogenous in the crime equation. In order to account for this possibility we use four strategies. First, we include a rich set of individual-level characteristics related to both crime and enrollment decisions, as well as county dummies. Many of these variables (e.g., family crime, certainty of punishment, number of children) are not commonly available, and thus would typically end up included in the error terms.

Second, we include the change in the number of schools per student (by county and year), as a measure of the change in the cost of attending school within each location, in the enrollment choice equation but not in the crime equation. The idea is that a higher concentration of schools per student should make it easier (less costly) to attend school. By using the number of schools per student as an exclusion restriction, it can work as a source of exogenous variation that aids in identification of the effect of enrollment on crime.[24]

Third, we factor analyze the residuals by taking advantage of some of the unique features of our data. As discussed earlier, one key advantage of our data is that it contains measures of both the cognitive and social/emotional skills of each individual, both of which may be important in driving both enrollment and crime decisions. Using these test measures, we first estimate a correlated factors model to isolate estimates of cognitive and social/emotional skills (see Section 3.1 for a description of the factor model we employ). We then include these measures of skills as regressors in our model, by decomposing the errors in equations (1) and (2) as follows:

$$\eta^e_{i,t} = \delta^{e,cog}\bar{\theta}^{cog}_i + \delta^{e,emo}\bar{\theta}^{emo}_i + \varepsilon^e_{i,t}$$

(crime to enrollment). The results are very similar. In Section 5 we discuss how this assumption affects the short-run and long-run impacts on enrollment and crime decisions through simulations of our model.

[23] We keep the conditioning on the remaining variables implicit to ease notation.

[24] We also tried estimating the model using both 2-year and 4-year college state-level tuition as an exclusion restriction in the enrollment equation, using tuition data from the Washington Higher Education Coordinating Board (HECB). The results were very similar.

$$\eta_{i,t}^c = \delta^{c,cog}\bar{\theta}_i^{cog} + \delta^{c,emo}\bar{\theta}_i^{emo} + \varepsilon_{i,t}^c,$$

where $\bar{\theta}_i^{cog}$ and $\bar{\theta}_i^{emo}$ are our estimates of cognitive and social/emotional skills, respectively.

Finally, while we assume that $\varepsilon_{i,t}^e$ and $\varepsilon_{i,t}^c$ are i.i.d. across individuals and over time, we allow them to be correlated with each other. The fact that we are able to observe a wealth of individual characteristics, which are highly persistent (or fixed) over time, as well as control for unobserved abilities through our factor estimates, allows us to pull components out of the error term that would otherwise generate correlation in the errors over time. In particular, we assume that the errors are jointly normally distributed and estimate the model using a bivariate probit.

The full model that we estimate is then a bivariate probit on $e_{i,t}$ and $c_{i,t}$ where

$$\begin{aligned} e_{i,t} &= \begin{cases} 1 & \text{if } v_{i,t}^e > 0 \\ 0 & \text{otherwise} \end{cases} \\ c_{i,t} &= \begin{cases} 1 & \text{if } v_{i,t}^c > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned},$$

where the latent variables $v_{i,t}^c$ and $v_{i,t}^e$ are given by

$$\begin{aligned} v_{i,t}^e &= z_{i,t}^e\beta^e + e_{i,t-1}\kappa^e + yc_{i,t}\lambda^e + ye_{i,t}\alpha^e + \delta^{e,cog}\bar{\theta}_i^{cog} + \delta^{e,emo}\bar{\theta}_i^{emo} + \varepsilon_{i,t}^e, \\ v_{i,t}^c &= z_{i,t}^c\beta^c + c_{i,t-1}\pi^c + e_{i,t}\gamma^c + yc_{i,t}\lambda^c + ye_{i,t}\alpha^c + \delta^{c,cog}\bar{\theta}_i^{cog} + \delta^{c,emo}\bar{\theta}_i^{emo} + \varepsilon_{i,t}^c, \end{aligned} \quad (3)$$

and where

$$\begin{pmatrix} \varepsilon_{i,t}^e \\ \varepsilon_{i,t}^c \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

3.1 Factor Model for Abilities

Let t_i and T_i denote the first and last ages for which individual i is observed in the data. Let M_{j,i,t_i}^{cog} denote one of $j = 1, \ldots, J$ cognitive measurements, where the t_i in the subscript denotes that the cognitive tests were given only in the baseline survey. We use 7 elements of a battery of tests that were taken by participants in the first wave of the survey. There are five continuous measures: the WASI Matrix Reasoning and Vocabulary scores and the three Stroop scores (Color, Word and Color/Word); and two Trail-Making scores which are measured on an ordered discrete scale.

We also include $k = 1, \ldots, K$ tests of social/emotional skills that are repeatedly measured in each survey, which we denote by $M^{cog}_{k,i,t}$. We employ three WAI scores: Impulse Control, Suppression of Aggression, and Consideration of Others; as well as three elements of the PSMI: Self Reliance, Identity, and Work Orientation.

For the case of the continuous measures, we write a linear model

$$\begin{aligned} M^{cog}_{j,i,t_i} &= x_{i,t_i}\beta^{cog}_j + \theta^{cog}_i \delta^{cog}_{j,t_i} + \xi^{cog}_{j,i,t_i}, \\ M^{emo}_{k,i,t} &= x_{i,t}\beta^{emo}_{k,t} + \theta^{emo}_i \delta^{emo}_{k,t} + \xi^{emo}_{k,i,t}. \end{aligned} \tag{4}$$

For the discrete Trail-Making measures that take L_j values, we let $\psi_{j,\ell-1} < \psi_{j,\ell},\ \ell = 1, ..., L_j$ with $\psi_{j,0} = -\infty, \psi_{j,L_j} = \infty$; and write an ordered model such that

$$\text{If } M^{cog}_{j,i,t_i} = \ell \quad \Rightarrow \psi_{j,\ell-1} < x_{i,t_i}\beta^{cog}_j + \theta^{cog}_i \delta^{cog}_{j,t_i} + \xi^{cog}_{j,i,t_i} \leq \psi_{j,\ell}. \tag{5}$$

$\theta^{cog}_i, \theta^{emo}_i$ denote cognitive and social/emotional abilities respectively, $\delta^{cog}_{j,t}, \delta^{emo}_{k,t}$ denote loadings that measure the effect of these skills, and the "uniquenesses" $\left\{\xi^{cog}_{j,i,t_i}\right\}_{j=1}^{J}, \left\{\left\{\xi^{emo}_{k,i,t}\right\}_{t=t_i}^{T_i}\right\}_{k=1}^{K}$ capture other determinants of the test scores like measurement error. While we assume that θ^{cog}_i and θ^{emo}_i are independent of the uniquenesses, we allow them to be correlated with each other. Identification of the factor model follows from the analysis in Carneiro, Hansen, and Heckman (2003) and Cooley, Navarro, and Takahashi (2015). Having obtained estimates of the parameters of the factor model, we then predict the most likely values for $\theta^{cog}_i, \theta^{emo}_i$ given the data we observe for each individual i.[25] These are the $\bar{\theta}^{cog}_i, \bar{\theta}^{emo}_i$ we use in equations (3).

4 Results

Before getting to the main results from our model, we first present the results from our factor analysis in which we project our measurements of skills onto two factors, one related to cognitive skills, and one related to social/emotional skills.

4.1 Factor Analysis

The results from the estimation of the factor model are presented in Tables 3-4 and Figure 4. We chose the following normalizations. The factor representing cognitive skills is normalized to have a loading

[25] See Appendix B for details on the estimation of the factor model as well as on prediction.

of one in the Matrix Reasoning WASI test score, while for the factor representing social/emotional skills the loading is normalized to one in the first period WAI Impulse Control measure. Besides being required for identification, these normalizations aid in the interpretation of the factors. Hence, the factor representing cognitive skills is such that an increase of one standard deviation in cognitive skills leads to an increase of one standard deviation in the Matrix Reasoning WASI test, and similarly for the social/emotional factor.

While we only allow the cognitive factor to affect cognitive measures and the social/emotional factor to affect social/emotional measures, we allow the two factors to be correlated. Our estimates show that there is more variance in social/emotional skills (0.19) than in cognitive skills (0.08), and the skills are positively correlated with a correlation coefficient around 0.23.

In Figure 4 we present a variance decomposition that allows us to get an idea of how important it is to account for measurement error (i.e., the uniqueness) when employing these measures. That is, we decompose the variance of the unobservable component of each measurement into the proportion of the variance coming from the skill (i.e., the factor) and the proportion contributed by the uniqueness.[26]

In Tables 3 and 4 we present the estimated parameters of the factor model for the measurement system. There are two interpretations for the coefficients on the covariates included in the factor model (e.g., gender, race). On the one hand, the coefficients can be interpreted as measuring differences in test scores that are unrelated to skills. For example, under this interpretation, the distribution of skills for men and women is the same, and hence the coefficient on the WASI Matrix Reasoning test of -0.023 in Table 3 would be interpreted as indicating that, on average, females perform worse on this test than a male of equivalent skills. On the other hand, they can be viewed as capturing differences in both test-taking and underlying skills.[27] Under this interpretation the coefficient on female reflects a combination of differences in skills and test-taking ability. Without further restrictions we cannot disentangle these two interpretations. Since we also include these variables in the crime and enrollment equations, this implies that our estimates of the coefficients on these variables in the crime and enrollment equations could be interpreted as reflecting combinations of direct effects and indirect effects via differences in skills. It does not, however, affect the interpretation of the other model parameters or of the simulations in Section 5.

As can be seen from Table 3, having more cognitive skills is related with having “better” scores in all of the cognitive measures we use. The negative sign for the Trail-Making scores is consistent with the way the scores are recorded where a larger score reflects cognitive impairment. As Figure 4 shows,

[26]In order to avoid having a graph for each age, we use the age-averaged proportions in our calculations.

[27]A third possible interpretation is that the coefficients reflect only differences in underlying skills. This interpretation imposes strong restrictions on the sign and magnitude differences across tests that are inconsistent with our estimates.

our measure of cognitive skills is more related to the Stroop measures of cognitive dysfunction than to the WASI-IQ and Trail-Making measures. However, even for the Stroop measures, cognitive skills can only explain at most 62% of the unobserved variance.

As documented in Table 4, for the case of social/emotional scores, more social/emotional skills lead to higher scores for all the social/emotional measures we include. There is also a general pattern consistent with maturation effects, in which the mean scores get better over time (i.e., the constant terms for each period in the equations) and social/emotional skills become a stronger determinant of the scores on the tests (i.e., the loadings). Social/emotional skills explain around 30% of the variance for all measures, except for the WAI-Consideration of Others where it essentially has no explanatory power. This result suggests that our measure of social/emotional skills is more related to individual discipline and control than to attitudes towards other people.

4.2 Baseline Model

We now present the results from our baseline specification. In Section 4.3, we consider several alternative specifications to evaluate the robustness of our results. In our baseline specification, in order to control for unobserved heterogeneity across individuals, we include our estimated cognitive and social/emotional skill estimates as regressors.[28] The results from the baseline bivariate probit are listed in column 1 of Table 5, where we report the average marginal effects of each covariate. We focus on the results for overall crime and discuss the results for the separate crime categories only when the results vary significantly by type of crime.[29] The results for drug-related, violent, and property crime separately are contained in the online appendix in Tables O1-O3.

We find that being in Maricopa County (Phoenix), compared to Philadelphia County, is associated with a higher probability of enrollment in school and a higher probability of committing crime. Blacks are less likely to engage in criminal activities and more likely to attend school compared to Whites. At the same time, Hispanics are less likely both to commit crime and to enroll in education than Whites, although the differences based on ethnicity are small and not precisely estimated. Females are more likely to attend school (5.8%-points) and less likely to commit crime (10.1%-points).

Consistent with what one would expect, having a "non-intact" family, is associated with lower enrollment rates and higher crime rates. Age is negatively associated with enrollment and crime. The result for enrollment is not surprising given that this dataset covers people between the ages of 14 and

[28] As a robustness check, in Section 4.3.5 we use the set of measurements used to infer the skills as regressors directly.

[29] Note that our results for overall crime should not be interpreted as an average across the crime categories, as the overall crime category pools all crimes together. However, we find that for most of our results, the overall crime estimates are consistent with the separate crime categories: violent, property, and drug.

26. The finding that crime also decreases with age is consistent with the broader empirical literature on the life-cycle of crime (Farrington, 1986; Hirschi and Gottfredson, 1983).[30]

Not surprisingly, the effect of the perceived risk of punishment has no effect on education and has a negative effect on crime, suggesting fairly strong deterrent effects of punishment: a 10% increase in the perceived probability of being caught generates a 2.2%-point decrease in the probability of committing crime, which is equivalent to a reduction in the crime rate of about 4%.[31] Each child an individual has decreases the probability of enrollment by about 1.8%-points, but has no effect on crime. Having family members involved in crime has a large positive effect on crime (14.9%-points), suggesting that the family environment plays an important role in determining criminal behavior. Perhaps a bit surprisingly, drug use has only a very small negative effect on enrollment decisions (0.1%-points). It has a large positive effect, however, on overall crime (22.4%-points).[32]

We also include the unemployment rate to control for local employment conditions. An increase in the unemployment rate by one percentage point leads to an increase in the probability of enrollment of 2.1%-points, or 4%. The effect of unemployment on crime is also positive but smaller in magnitude (1%-point or 2%). These results suggest that criminal youth respond to worsening economic conditions by staying in school and, to a lesser extent, increasing criminal activity. Our results are consistent with those of Betts and McFarland (1995) and Dellas and Sakellaris (2003) who find that a one percentage point increase in the unemployment rate leads to an increase in enrollment in college by about 4%. With regards to crime, Raphael and Winter-Ebmer (2001) and Gould, Weinberg, and Mustard (2002) estimate that a one percentage point increase in the unemployment rate generates an increase in crime of between 1 and 5%.

We also included a measure called the Future Outlook Inventory, which measures the degree of future consideration and planning, and proxies for the individual's discount factor. Low discount factors is one potential cause of criminal activity (Davis, 1988; Mastrobuoni and Rivers, 2015), as people who care less about the future may be less deterred by the future consequences of their actions. Similarly, high discount factors are associated with higher investment rates (Chen, 2013; O'Donoghue and Rabin, 1999), such as investing in education. Our results are consistent with this, as the sign on the effect of Future Outlook Inventory is negative for crime and positive for education.

As discussed in Section 4.1, higher values of our estimates of cognitive and social/emotional skills

[30] Drug crime does not seem to decrease with age. Combined with the statistic from Table 1 that shows that people start committing drug crimes at much later ages, this suggests that the age profile for drug crime is different compared to violent and property crime (Sampson and Laub, 2003; Farrington, 1986; Wilson and Herrnstein, 1985).

[31] These findings are in line with Lochner (2007), who finds that a 10% increase in the perceived probability of arrest reduces criminal participation in major thefts by about 3% and in auto theft by more than 8%.

[32] This result is not solely driven by the effect on drug-related crime. The effects on violent crime (15.9%-points) and property crime (14.4%-points) are also quite large.

are associated with better performance on the tests. Therefore, we should expect them to be positively associated with education and negatively associated with crime. We find that higher cognitive skills increase the likelihood of enrollment and higher social/emotional skills lead to lower crime rates. The results imply that a one standard deviation increase in social/emotional skills leads to a decrease in the probability of crime of 3.5%-points. Also, a one standard deviation increase in cognitive skills leads to an increased probability of enrollment of 1.0%-points, although it is not precisely estimated. The effects of cognitive skills on crime and social/emotional skills on education are both small and imprecisely estimated.

Initially we expected these effects to be larger (see e.g., Cawley, Heckman, and Vytlacil, 2001; Heckman, Stixrud, and Urzua, 2006; Murnane, Willett, and Levy, 1995). However, there are several reasons for why we would find more moderate effects. First, we are able to control for a very rich set of observables, many of which are not commonly available in other datasets. In the absence of data on these individual characteristics, their effects will be conflated with the effects of skills, biasing estimates of their effects by causing the skill measures to have to explain more of the variation in enrollment and crime decisions. Second, because the sample consists of serious juvenile offenders only, the distributions of both types of skills are compressed relative to the population at large. As a result, a one standard deviation change is not particularly large in our data.

In addition to controlling for many sources of individual heterogeneity, we also examine the effect of contemporaneous education on crime. In order to account for the possibility that enrollment is endogenous, we include the change in the number of schools per student as an exclusion restriction in the enrollment equation, but not in the crime equation. We find that more schools per student is strongly positively related to enrollment, consistent with the idea that a higher concentration of schools makes it less costly to attend school.

We find that enrollment leads to an increase in overall crime rates (8.8%-points).[33] The effect varies by the type of crime though. For property crime, we find weak evidence that enrolling in school decreases crime, with an average marginal effect of 2.3%-points that is not precisely estimated. This is consistent with the incapacitation effect found by Jacob and Lefgren (2003); Luallen (2006); and Anderson (2014), although our effect is smaller in magnitude.

For violent and drug-related crime, we find the opposite effect: enrollment leads to an increase in crime rates (10.4%-points for violent and 7.7%-points for drug-related). This suggests the presence of

[33] One potential concern with this result is that not being enrolled is a proxy for being incarcerated, and therefore this estimate captures the incapacitation effect of prison. This is unlikely here, as in our data the relationship between enrollment and incarceration goes in the other direction as they are positively correlated. See also Section 4.3 in which we discuss the effects of being in jail in more detail.

positive complementarities between school and drug/violent crime. This is consistent with the concentration story of Jacob and Lefgren for violent crime–that an increased density of young people leads to more violent interactions. For drug-related crime, one explanation is that the primary buyers of drugs sold by juveniles are other juveniles, and thus attending school allows the sellers of drugs to be closer to their clients.

The last row of Table 5 reports the correlation in errors of the crime and enrollment equations. The estimate of -0.142 indicates that the remaining unobserved drivers of crime and education decisions are negatively correlated with each other, although the correlation is not precisely estimated. As we show in the next section, failing to account for this negative correlation leads to a downward bias in the estimate of the contemporaneous effect of enrollment on crime.

Finally, we allow for previous crime and education decisions to affect current decisions in two ways. First, we allow the lagged decisions to affect the current ones.[34] This captures state dependence, or inertia, in these decisions. Second, we also allow the total accumulated experience (measured in years) to affect decisions. The rationale for this is that human and criminal capital accumulated through previous educational or criminal experience could affect the returns to both school and crime (Lochner, 2004; Nagin and Paternoster, 1991; Nagin, Farrington, and Moffitt, 1995; Imai, Katayama, and Krishna, 2006; Merlo and Wolpin, 2015; Loughran et al., 2013).

We find strong evidence of state dependence in both the education and crime decisions (Brame et al., 2005). Enrolling in school the previous period increases the probability of enrolling in the current period by 18.9%-points. Participating in crime in the previous period increases the probability of crime by 15.8%-points. We also find some evidence of returns to experience, although the effects are smaller. The signs of the results are as expected. An additional year of education is positively associated with enrollment decisions and negatively associated with crime, but the effects are small and not statistically significant. The effect of criminal experience on crime is positive: an extra year of criminal experience increases the probability of crime by 2.0%-points. The effect on education is negative, with an extra year of crime associated with a decrease in the probability of enrollment by 0.7%-points.

Overall our estimates suggest that there are important dynamics in both the crime and education decisions. While both matter, the effects of state dependence are much larger than the returns to experience. This distinction is relevant for policy, as understanding how the pattern of previous decisions drives current decisions is important for determining how and when to attempt intervention. We discuss

[34]For simplicity, in our baseline model we allow for lagged crime to affect current crime and lagged education to affect current education, but do not allow for lagged cross-equation effects. We also tried estimating a version allowing for these effects. The coefficients on these additional terms were small and statistically insignificant. The other estimates were virtually unchanged, with the exception of the effect of contemporaneous enrollment on crime, which increased slightly.

this more in Section 5 when we illustrate these effects with various simulations based on our model.

4.2.1 The Effect of Education on Crime

Enrollment Our results regarding the effect of contemporaneous enrollment on crime are generally consistent with the results of Jacob and Lefgren (2003) and Luallen (2006), who examine the effect of short-duration shocks to school attendance. However, the more direct comparison is probably to Anderson (2014), as he examines the effect of compulsory schooling laws designed to keep youth in school for additional years. Anderson (2014) finds that compulsory schooling laws decrease violent, property, and drug crime (although the results for drug crime are not precisely estimated), consistent with crime-reducing effects of enrollment.

One explanation for the differences in our findings is based on the crime measures employed. Anderson (2014) uses arrests, as opposed to self-reports, as the crime measure. His measure presumably contains a higher proportion of more severe crimes, and contains proportionately fewer minor offenses such as fighting and drug dealing, as these are less likely to result in an arrest.[35] As discussed in Anderson (2014), it may be that enrollment leads to an increase in these minor crimes via the concentration story of Jacob and Lefgren, but leads to a decrease in more serious offenses. When we exclude fighting from our measure of crime, we find that the contemporaneous effect of enrollment on crime drops by half and is no longer statistically significant. It shrinks further if we exclude drug offenses, although the point estimate remains positive. Overall it appears that heterogeneity in the composition of crime severity captured by arrests versus self-reported crime data may be driving some of the differences in our results.

Another difference in our crime measures is that we analyze the extensive margin of crime, whereas aggregate crime measures capture the intensive margin as well. While enrollment may lead to a reduction in the intensive margin, it may not drive it to zero, particularly for our sample of serious offenders, resulting in smaller estimated effects on the extensive margin. In order to examine this, we tried re-estimating our model using continuous measures of crime intensity. The estimated effects of enrollment were negative overall, but small and statistically imprecise, suggesting at most a small role for the intensive margin as an explanation for the differences in our findings. Finally, as discussed in Durlauf, Navarro, and Rivers (2008, 2010), crime regressions based on aggregate data (which is the case for all three papers discussed above) can yield very different results than those based on individual-level data.

[35] Luallen (2006) also employs arrests as the outcome, while Jacob and Lefgren (2003) use reported incidents.

Educational Attainment Our baseline results suggest that years of schooling have no significant effect on crime. An additional year of education decreases the probability of crime by 0.3%-points (0.5%), and the effect is not statistically significant. Nevertheless, several studies suggest that educational attainment is an important determinant of adult crime. Starting with the seminal work of Lochner and Moretti (2004), several studies employ changes in compulsory schooling laws over time in order to control for the potential endogeneity of education decisions, using a variety of crime outcome measures. Lochner and Moretti (2004) find that a one-year increase in schooling leads to increases in annual arrest and incarceration rates of approximately 18% and 11-25%, using data from the US Census and Uniform Crime Reports (UCR). Hjalmarsson, Holmlund, and Lindquist (2015) use Swedish data and find that an additional year of education reduces the probability of ever being convicted by 6.7% and ever incarcerated by 15%. Using data from England and Wales, Machin, Marie, and Vujić (2011) report that a 10% increase in age-left-school leads to a 2.1% decrease in annual convictions, which translates to a 1.3% decrease for an additional year of schooling. These results are based on measures of arrests, convictions or incarcerations (and at different time intervals), whereas our results are based on annual self-reported crime measures. One possible reason for why we do not find strong evidence of an effect of educational attainment on crime may be due to the crime measures being employed.

The two sets of results most closely related to ours are Lochner and Moretti (2004) and Merlo and Wolpin (2015), which both employ annual self-reported crime data. Using data on young men in the NLSY79, Lochner and Moretti (2004) find that an additional year of school reduces participation in crime by 2 to 3%-points (10%).[36] Merlo and Wolpin (2015) estimate a multinomial discrete-choice VAR model of crime, education, and employment on a sample of black males from the NLSY97, which allows for lagged effects (state dependence), but not experience directly. They find that not attending school at age 16 increases an individual's crime rate by around 13%.

It is possible that, for our selected sample of serious criminal offenders, educational attainment does not play a relevant role in deterring crime. For example, individuals in our data may benefit little in terms of labor market opportunities from additional schooling, given their existing criminal history (Waldfogel, 1994 and Kling, 2006). It could also be that after controlling for a richer set of observables, in particular some that are usually not available in other datasets, educational attainment is largely unimportant.[37] It may also be the case that the quality of the education received by individuals in our sample is lower, for example due to some of the education being received while in a locked facility.

[36] Unlike their analysis for arrests and incarceration, the NLSY79 results of Lochner and Moretti (2004) do not instrument for educational attainment.

[37] Tauchen, Witte, and Griesinger (1994) and Witte and Tauchen (1994) find little evidence of an effect of educational attainment on crime after controlling for previous criminal activity and several individual characteristics.

Finally, it is also possible that the relevant margin for education is high school graduation, not years of education *per se* (Lochner, 2011b).

In order to explore these alternative explanations, we estimate several additional specifications of our model. We re-estimate our model using a reduced set of controls, specifically only location, non-intact family, age, the unemployment rate, and IQ. We also drop lagged effects and criminal experience and assume that errors across equations are uncorrelated. This roughly corresponds to what the previous literature includes. For both the full and reduced set of controls, we estimate the model on the full sample, the sample of males only (as most of the literature focuses on males), and the sample of males using an alternative measure of years of education for all ages and for those at least 18 years of age. Our alternative measure of years of education does not include years of education obtained while in a locked residential facility. The motivation is that education obtained while incarcerated may have smaller crime-reducing effects. The results are reported in Table 6.

As is illustrated in the first set of results, once we drop lags, criminal experience, and the additional controls, our results are closer to those in the literature.[38] The results suggest that an additional year of education is associated with a decrease in the probability of crime ranging from 2.0 to 2.8%-points (4-5%), as we restrict the sample to correspond more closely to what the literature has used. In contrast, the results for the full set of controls are much smaller in magnitude than those in the literature. There is also a small increase in the absolute value of the effect when we employ our alternative measure of education, consistent with quality differences in education obtained while incarcerated.

Unfortunately we do not directly observe high school graduation. As a proxy we estimated a specification with a dummy for 12 years of educational attainment. We do not report these results, as the coefficient on the dummy for 12 years of schooling on crime was very small and insignificant, and the coefficients on the other variables changed very little.

4.3 Alternative Specifications

In this section we present results from two sets of alternative specifications to our baseline model that are designed to illustrate how our modeling choices affect the estimates. In columns 2-6 of Table 5 we include simple variants to our baseline identification strategy. In particular, we estimate versions of the model in which we incorporate only a limited set of control variables; do not allow for the crime and education equation errors to be correlated (independent probits instead of a bivariate probit);

[38] We also compare to Merlo and Wolpin (2015) by dropping our rich set of controls to more closely match their setup and simulating long-run effects. Our results for the effect of prior education on crime are larger and statistically significant, but still smaller, compared to Merlo and Wolpin (2015) (-2% vs. -13%). We do, however, find very similar long-run effects of prior crime on crime.

do not allow for dynamics; do not include the number of schools per student as an exclusion in the enrollment equation; and use the direct measures of cognitive and social/emotional skills, as opposed to our estimates of the underlying skills from the factor model.

The objective for the second set of results is to provide some additional robustness checks to the baseline model.[39] We show that our results are robust to alternative ways to treat decisions while in jail; excluding drug use as a control; alternative definitions of enrollment; allowing the effects of prior crime and education decisions, as well as contemporaneous enrollment, to vary by age; alternative specifications for criminal experience; and switching the contemporaneous effect from crime to enrollment.

4.3.1 Controls

A key benefit of our data is that we are able to control for a rich set of observable (criminal involvement of the family, expected probability of punishment, and degree of future consideration, among others) and typically unobservable (cognitive and social/emotional skills) sources of individual heterogeneity, that are not commonly available in other datasets. Since most of these variables are highly persistent over time (or fixed), failing to control for them could lead to estimates of the dynamic effects that are biased upwards in absolute value. In order to see the possible extent of this bias, we estimate a version of our model in which we include only a sparse set of individual characteristics and the local unemployment rate. The results are reported in column 2 of Table 5. Consistent with our hypothesis, we find that the estimated effects of lagged criminal and educational decisions are inflated, particularly their effects on crime. The returns to criminal experience on crime almost double from 2.0 to 3.9%-points, and the effect of lagged crime increases by roughly 50% from 15.8 to 23.5%-points. The effects of educational experience on both crime and enrollment also increase and become statistically significant (from -0.3 to -1.4%-points and from 0.6 to 1.1%-points, respectively).

4.3.2 Uncorrelated Errors

In order to determine the importance of allowing the errors in the crime and education decisions to be correlated, we re-estimate the model using separate probits for the two equations, rather than a bivariate probit model. The estimated effects are very similar between the two models, with the exception of the effect of current enrollment on crime, which drops from 8.8 to 2.5%-points. In the bivariate probit model, the errors are estimated to be negatively correlated with each other. When we assume that they are independent (and therefore uncorrelated), the model has to decrease the direct effect of current

[39] We present these results in Tables A1-A4 in Appendix A and in the online appendix Tables O4-O6 for the crime-specific estimates.

enrollment on crime to account for this and fit the data, leading to a substantial underestimate of the causal effect of enrollment on crime.

4.3.3 No Dynamics

The intuition for the effect of not including dynamics in the model is similar to that for not including covariates. To the extent that there are important dynamic relationships, excluding them from the model will lead to the magnification of the effects for the other included variables. In column 4 of Table 5, this is exactly what we see. When we do not allow accumulated experience and lagged decisions to enter the model, the effects of the individual heterogeneity increase in absolute value, overstating their true contribution. For example, the effect of drug use on crime increases from 22.4 to 26.7%-points. The average marginal effect of social/emotional skills on crime also increases in magnitude from -8.0 to -12.7%-points. For the same reason, this also changes the estimates of the contemporaneous effect of enrollment on crime, more than doubling the estimated effect from 8.8 to 20.2%-points. This highlights the importance of controlling for the dynamics in the crime and education decisions. Even when the object of interest is not dynamic, failing to account for dynamics causes biased estimates of other relationships, including the contemporaneous effects.

4.3.4 Not Instrumenting

As we discuss above in Section 3, in order to address the potential endogeneity of enrollment in the crime equation, we introduce an exclusion restriction by adding the change in the number of schools per person in the enrollment equation. In column 5 of Table 5 we present results in which we do not include this, in order to illustrate its effect on our estimates. The primary concern was that failing to appropriately control for endogeneity would lead to a biased estimate of the effect of enrollment on crime, which could in turn generate bias in the other estimates as well. We find that by not including this variable, the estimate for contemporaneous enrollment drops from 8.8 to 6.5%-points. The difference is consistent with the expected bias given the negative correlation of the errors. This result demonstrates that there is some bias that this exclusion restriction is correcting for. However, the bias is not particularly large, which is likely due to the fact that our data allow us to control for many sources of observed and unobserved heterogeneity that would otherwise generate further correlation in the errors of the crime and enrollment decisions, and exacerbate the endogeneity problem.

4.3.5 Cognitive and Social/Emotional Skills

We also estimate a specification in which we replace our estimates of skills with the measures used to infer them. This allows us to investigate whether our results are sensitive to our use of the estimated cognitive and social/emotional skills, and also to better understand how cognitive and social/emotional skills contribute to enrollment and crime decisions. As can be seen in column 6 of Table 5, the estimates on the other variables are very similar to the baseline estimates, illustrating that our factor-model-generated measures are effective summaries of these skills.

A somewhat surprising result is that the two measures that generate the IQ score (Matrix Reasoning and Vocabulary) have no effect on enrollment decisions. The point estimates are very small and insignificant. Given that cognitive skills are viewed as one of the primary drivers of education decisions in the literature, this is particularly surprising. One explanation for our finding is that the IQ distribution in our dataset is substantially shifted to the left, compared to the general population. Themedian raw IQ score is only 85 in our data, with only about 10% scoring above the population average of 100. It may be that in this range of IQ scores, marginal increases in IQ do not have significant effects on the value of education or on the cost of completing education. In contrast, one of the measures of cognitive impairment does seem to be related to education decisions. The Trail-Making B test, which involves the sequencing of number and letters is negatively associated with enrollment. So while IQ scores do not seem to be significant drivers of enrollment decisions, there is some evidence that cognitive impairment does. In particular the Trail-Making B test seems to be the cause of the positive correlation between cognitive skills and enrollment in the baseline specification.

Consistent with the baseline estimates, the tests for cognitive skills are generally uncorrelated with crime decisions. The sole exception is for property crime, in which there seems to be evidence of positive returns to cognitive skills.

We have six measures of social/emotional skills. These measures have a consistent negative effect on crime (most of which are statistically significant), with the exception of the PSMI-Self-Reliance measure, which has a positive sign. These results are consistent with the literature, which finds that a lack of social/emotional skills can be an important driver of criminal activity. For example, Gottfredson and Hirschi (1990) suggest that the inability to exercise self-control (measured as WAI-Impulse Control and WAI-Suppression of Aggression in our data) can explain a large part of criminal behavior. The fact that self-reliance, which is viewed as a positive trait, is associated with a higher probability of committing crime, suggests that some social/emotional skills may be beneficial for both legitimate and illicit activities.

Overall the social/emotional measures have small and insignificant effects on enrollment, consistent with our baseline results. However, two components of the PSMI appear to be important for schooling decisions. PSMI-Identity has a positive effect on enrollment, which makes sense since this measures self-esteem and consideration of life goals. Somewhat surprisingly, PSMI-Work Orientation has a negative effect on enrollment.

4.3.6 Modeling Choices While in Jail

In our dataset we can distinguish whether individuals attended a community school only, an institutional school only, both community and institutional schools, or none, during each recall period. The decision and the incentives to attend institutional schools when an individual is incarcerated may be different from enrolling in a community-based school when the individual is free. Unfortunately, we cannot distinguish between a person who was free during some portion of the recall period, and chose not to go to a community school, and a person who did not have the choice at all because he was incarcerated throughout the whole period. Furthermore, we cannot observe whether crime choices during the recall period were made while free or incarcerated. In our baseline specification we drop observations in which an individual attended only an institutional school in a given year.

In order to determine if our results are sensitive to this choice,[40] we estimate three other model specifications. In the first, we set enrollment to zero if an individual did not attend a community school (i.e., attended an institutional school only, or attended no school). In the second specification, we add a variable to the model that is an indicator for whether the individual was incarcerated at the time of the interview, to allow for being in jail to affect choices. Finally, we add the indicator for jail interacted with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail. The results of the three specifications are reported in Table A1.

In the first specification, the marginal effects for female, punishment, family crime, and drug use increase in absolute value in the enrollment equation. This is likely to due the fact that these are strong predictors of crime. When we assume that people who attend only institutional schools decided not to attend community school (instead of excluding those observations from the likelihood), we are adding observations in which people are incarcerated and not attending school. Therefore any variables which predict that people are more (less) likely to commit crime, will predict that these people are more (less) likely to be incarcerated, and therefore less (more) likely to enroll in school. This is exactly the pattern

[40] See Piquero, Schubert, and Brame (2014), who find that controlling for time spent in prison is important for interpreting time series patterns in offending.

that we see for female, punishment, and family crime.

While drug use is also a strong predictor of crime, the explanation above would cause the effect of drug use on crime to become more negative (drugs cause more crime, more incarceration, and thus less school). However, we observe the opposite. The most likely explanation here is that it is more difficult to obtain and use drugs while in jail, so adding these observations (in jail and not attending school) generates a positive correlation between drug use and enrollment.

The effect of years of education on enrollment also increases and becomes statistically significant, although the effect is still not that large (2.3%-points). One possible explanation is that people who are incarcerated have few years of schooling, so by adding these observations (few years of education and not attending school) we are reinforcing the positive correlation between experience and education choice. We also observe a small decrease in the effect of contemporaneous enrollment on crime. This is also likely due to the addition of observations for individuals who were both not attending school and incarcerated (and therefore likely to have committed a crime in that period).

When we condition on being in jail, the effect of enrollment on crime decreases slightly, but overall the results are quite similar to those in the baseline. When we interact the dummy for being in jail with our measures of education and crime, we find that our main results are largely unchanged compared to the specification with just the dummy for jail. The only difference is that we observe some evidence that the returns to previous educational and crime choices are lower while in jail. The interaction between jail and lagged enrollment and educational experience in the enrollment equation are negative, and lagged crime interacted with jail is also negative.

Overall our results with respect to modeling the choices while in jail suggest that our baseline results are quite robust to alternative modeling decisions. While some of the results related to individual characteristics are affected in some cases, our main results about the contemporaneous and dynamic relationships between crime and education are largely unchanged.

4.3.7 Drug Use

Another potential concern relates to the fact that drug use is a choice rather than an exogenous variable, which may bias some of our results. In particular one might think that education affects the propensity to use drugs, and that our finding that drug use has a strong positive effect on crime, and educational attainment does not, masks the indirect effect of education on crime via drug use. In order to check for this possibility, we estimated a specification of our model in which we drop drug use. The results are reported in column 1 of Table A2 in Appendix A. Dropping drug use does not change the effect of education on crime through educational attainment, suggesting that education does not have much effect on crime either directly or indirectly through drug use. On the other hand, dropping drug use increases the estimates of the effects of both skill measures on crime, by about 4%-points each. This suggests that skills have not only a direct effect (which is what we capture in the baseline estimates), but also an indirect effect through drug use. The results for the other coefficients are largely unchanged.

4.3.8 Defining Enrollment

In our baseline model we define an individual as enrolled in school if they are enrolled in school at the time of the interview, or if they were enrolled prior to coming to their detention facility. In order to determine if our results are sensitive to this, we re-estimate the model under an alternative definition of enrollment by defining enrollment as having attended school for at least nine months in the previous year. (We also adjust years of education and lagged enrollment accordingly).[41] The results are reported in column 2 of Table A2. Our main results are largely unchanged.

4.3.9 Age-Varying Coefficients

One potential concern with our baseline specification is that, if the effects of previous and contemporaneous education and crime decisions vary by age, then any estimated effects, particularly long-run effects, may be biased. In order to examine whether, and to what extent, this may be the case, we estimate a version of the model in which we allow the effects of accumulated experience, lagged decisions, and contemporaneous enrollment to vary by the age of the individual. In particular, we interact these variables with a dummy for whether the person is over 19 years old. In column 1 of Table A3 we find that the estimates vary slightly by age, but the differences are small. The largest change is in the effect of lagged enrollment on education, in which the marginal effect decreases with age from 22.9 to 17.0%-points, suggesting that the state dependence in educational decisions decreases slightly as individuals age, which is not surprising. Overall, the results seem to be consistent across age.

4.3.10 Criminal Experience

In the baseline survey we observe the age at which individuals first engage in crime, but we do not have a measure of accumulated criminal experience at the time of entry into the survey. In our baseline model we impute the accumulated years of crime using the procedure described in Section 2. Our estimates suggest a larger role for state dependence compared to returns to experience. One possible explanation

[41] We also estimated a version of the model in which we treated enrollment in months as a continuous outcome. Although the interpretation of the results is slightly different, the results were qualitatively similar to the results for defining the cutoff to be nine months.

for this result is that experience enters utility in a non-linear fashion, causing us to not fully capture its impact, whereas lagged crime is a dummy variable, and therefore already enters the model flexibly.

In columns 2 and 3 of Table A3, we allow experience to enter quadratically and as a piecewise-linear function of experience, allowing for different returns for 0-4, 5-9, and 10+ years of criminal experience. When we allow experience to enter quadratically we find that, consistent with the baseline, criminal experience has a small negative effect on enrollment and a positive and increasing effect on crime. However, we lose statistical significance on all of the associated parameters. In the second specification, the effect of years of crime on crime is similar to the baseline with no significant variation across the different experience categories.

Another concern is that our imputation procedure generates a noisy measure of criminal experience, making it more difficult to tease out the true returns to experience. In column 4 of Table A3, we use only the observed accumulated experience after entry, interacted with age of entry dummies, instead of our imputed measure.[42] Since observed criminal experience is likely to be positively correlated with the unobserved experience that occurs prior to entering the survey, we should expect that the coefficients will be inflated, as they will capture the effect of both the observed and unobserved experience. This upward bias in the coefficients is likely to be increasing in the age of entry into the survey, since the unobserved period is longer for people who entered the survey at an older age. This is consistent with our estimates. Furthermore, even if we ignore the bias in these coefficients, we still find a larger impact of lagged crime compared to criminal experience. Overall we conclude that our finding that state dependence has a stronger effect on crime than criminal experience is not driven by measurement or specification issues related to experience.

4.3.11 The Contemporaneous Effect of Crime on Education

In our baseline model, we estimate the contemporaneous effect of education on crime. As discussed above, we could have alternatively estimated the contemporaneous effect from crime to education.[43] In Table A4 we present estimates from this alternative specification. The results in the first column show that contemporaneous crime leads to an increase in enrollment of 9.7%-points, which is similar in magnitude to our estimate of the effect of enrollment on crime in our baseline specification. The correlation in errors of the crime and enrollment equations is negative, although not precisely estimated, as was the case in the baseline. The results for other coefficients are relatively unchanged. In column 2 we include lagged state arrest rates as an exclusion restriction in the crime equation (to serve as an instrument

[42] As discussed in footnote 17, our alternative procedure for accounting for unobserved criminal experience also gives us similar results.

[43] But not both. See our discussion in Section 3.

for contemporaneous crime in the enrollment equation).[44] The coefficient on contemporaneous crime increases slightly to about 12%-points, while the other coefficients remain largely unaffected.

5 Model Simulations

In this section we attempt to disentangle the roles of state dependence (i.e., lagged choices), criminal and human capital (i.e., accumulated years of crime and education), and heterogeneity both in terms of "observables" such as the perceived probability of punishment and "unobservables" such as skills, in driving the interactions between education and crime. Understanding the importance of each of these determinants is crucial, as the policy recommendations associated with them are quite different. For example, if state dependence is important and criminal activity is very persistent, then preventing someone from committing a crime at an early age will have important effects on future criminal activity as the persistence will tend to reduce crime even if nothing else is changed. Furthermore, if being enrolled in school has a large effect on whether one commits a crime or not, enrollment policies may be an important alternative to other incapacitation policies like incarceration. If, on the other hand, other determinants of crime (e.g., skills) are more important, then one should consider policies that foster these skills.[45]

For this purpose, we present two types of simulations based on our estimated baseline model. In the first case, we try to isolate the importance of dynamics by comparing the predicted paths of enrollment and crime decisions for two identical individuals (with median characteristics), who differ only along one dimension in the initial period (i.e., temporary differences). In particular, we simulate how these paths differ for an individual that commits a crime at age 15 from one that does not, and similarly for attending school at age 15. We do the same for two individuals with perceived probabilities of punishment that differ by 10%-points at age 15 and are equal in all subsequent periods. In the second set of simulations, we trace the dynamic effects of permanent differences in variables that measure heterogeneity, specifically differences in cognitive skills, social/emotional skills, and the perceived probability of punishment (i.e., a permanent 10%-point difference).[46]

[44] Data on state-level arrest rates was obtained from the FBI's Uniform Crime Reports.

[45] Cunha et al. (2006) provide evidence that very early periods are the most important for skill development. To the extent that education is still a key driver of skill development for the sample we study (adolescent and early-adult criminals), policies designed to promote enrollment in later years could provide additional crime-reducing benefits via skill formation.

[46] In our model we are assuming that skills are fixed over the age range we study. In this sense, our estimated effects of education on crime and crime on education are estimates of direct effects, holding skills constant. To the extent that education or crime also affect skill formation, there is an indirect effect captured by the skill channel. Ideally one would endogenize the process for skill formation in order to measure this channel directly. However, such a model would involve additional issues of simultaneity due to complicated feedback effects between enrollment/crime choices and skills. In addition, in our data some of the skill measures are only observed in the baseline survey, making it difficult to measure how skills evolve over time.

5.1 Dynamic Effects of Temporary Differences

We begin by simulating the differences from committing versus not committing a crime at age 15. Figure 5 shows that this has a very small effect on the probability of enrolling over time. The probability differs by 1.7%-points after 5 periods (from a baseline of 40%), and then it decreases as a consequence of aging since, after 10 years, almost no one in the data is enrolled anymore. Figure 5 shows that the effects on crime are much larger. Mechanically, the difference in the probability of committing a crime at age 15 is one. After one year, the probability of committing a crime is lower by 20%-points, from a baseline of 70%. This effect is almost entirely a consequence of state dependence (i.e., lagged crime). After that, the effect diminishes over time but, because of the decrease in criminal experience, it does not disappear. After 10 years, the person who did not commit a crime at age 15 is approximately 6%-points less likely to commit a crime.

Next, we analyze enrollment in school at age 15. In Figure 6 we can see that the effect of education on education is very similar to the effect of crime on crime. Mechanically the difference in the probability of being enrolled is one at age 15. As a consequence of state dependence, the probability is around 20%-points higher after a year. It decreases over time, reaching zero after 10 years. Its effect on crime is small but not insignificant (at least in the first years). Since enrollment has a positive contemporaneous effect on crime, as we can see in Figure 6, it increases the probability of crime by 8%-points initially. The effect rapidly decreases, and it reaches zero after 3 years. After that, it becomes slightly negative but very small as more and more human capital (i.e., years of education) gets accumulated.[47]

The third simulation we present, the effect associated with a 10%-point difference in the perceived probability of punishment at age 15, is shown in Figure 7. The effect on enrollment is negligible. Its effect on crime, on the other hand, is larger. At age 15, it reduces the probability of committing a crime by almost 2%-points. While the effect decreases rapidly, 10 years later there is a 0.1%-point lower probability of committing a crime.

Overall, the initial differences persist somewhat in the short run, and then decrease towards zero after several years. This is due to the fact that returns to experience are small relative to the effects of state dependence and individual heterogeneity. This implies that while policies based on temporary interventions will have only small effects on behavior many years after the policy (and thus may have

[47]As we mention in Section 3, we also estimated a version of the model in which there is a contemporaneous effect of crime on enrollment instead of an effect of enrollment on crime. In all of the simulations we describe in this section the long-run outcomes are very similar between the two model specifications. In a few cases, the short-run effects are different. In particular, for the case of the difference in committing a crime at age 15, in the alternative specification there is a short-run negative effect on enrollment that does not appear in our baseline model. Similarly, for the case of a difference in enrolling in school at age 15, there is no longer a short-run positive difference in crime. In both cases these differences diminish quickly. Figures for simulations from this alternative specification that are analogues to Figures 5-10 are located in Figures O1-O6 in the online appendix.

to be repeated to continue the effect), the potential gains to such policies are not insignificant. Given that crime is highly concentrated among young people, obtaining immediate and somewhat persistent reductions in crime has the potential to significantly affect overall crime rates.

5.2 Dynamic Effects of Permanent Differences

We next consider the effects that permanent differences in heterogeneity (while holding all other characteristics at their median values) may have on both the enrollment and crime probabilities. We begin by simulating paths of an individual with cognitive skills at the 25th percentile and comparing to one with skills at the 75th percentile in the data. While this may sound like a large difference, this is for individuals in our selected data where this distribution is much more compressed than in the overall population. For example, the 25th and 75th percentiles of the cognitive skill distribution are associated with IQ scores of 89 and 98 and scores of 39 and 48 on the Stroop Word test, respectively–a modest difference.[48] Figure 8 shows the effect on enrollment. Not surprisingly, higher cognitive skills are associated with a larger probability of being enrolled, but the magnitude of the difference is small: at most 3%-points (after five years). Cognitive skills are essentially not related to the probability of crime.

Figure 9 shows similar results for social/emotional skills. A movement from the 25th to 75th percentile for these skills is equivalent to a one-third of a standard deviation difference in impulse control, for example. As can be seen from the figures, the effect on enrollment is negligible. A different story arises when we look at the effect on criminal activity. The probability of committing a crime is lower by 3%-points for the individual with higher social/emotional skills at age 15, and the effect keeps growing over time. After 10 years the probability of committing a crime is reduced by 10%-points.

The final simulation is shown in Figure 10. In this case we simulate the paths based on a permanent 10%-point difference in the perceived probability of punishment. After five years the probability of enrollment is marginally larger, by less than 0.7%-points. The impact on crime is more significant. At age 15, the probability of crime is almost 2%-points lower for the individual with the higher perceived probability of punishment and the difference gets larger over time. After ten years it is almost 5%-points.

6 Conclusion

In this paper, we show that distinguishing between the potential sources of persistence in enrollment and crime decisions is important both in terms of generating a better understanding of what drives

[48]In order for a Word score to be considered "higher" or "lower" than another, a 10 point or greater score difference is required.

behavior, and for the purpose of designing policy. We find that individual heterogeneity is strongly related to criminal behavior. Many of these dimensions of heterogeneity go beyond what is typically measured in most datasets, such as attitudes about the future (future outlook inventory), drug use, family crime, and social/emotional skills. This illustrates the importance of controlling for a rich set of individual characteristics. Our results also help to identify which particular sources are most relevant for driving behavior. In particular, we find that social/emotional skills are important drivers of criminal behavior.

While we do not directly simulate potential policies designed to increase enrollment and/or decrease crime, our model simulations illustrate how policies targeted at altering individual heterogeneity (e.g., social/emotional skills) would drive changes in education and crime over time. We find, perhaps unsurprisingly, that permanent or long-run changes generate the largest effects. However, policies with temporary changes to individual behavior, such as keeping people out of crime for one period, can also have lasting effects. For example, a policy that prevents someone from committing a crime in a given year would generate an effect on crime in the following year of -18%-points. This implies that there is room for policies designed to shock individuals out of current bad decisions, and thus break the persistence caused by this state dependence. To the extent that these types of policies are easier to implement than permanent changes to individuals, their effect should not be dismissed. The reductions obtained are considerable and, at least in the case we model here, they are obtained during the ages in which criminal activities are at their peak.

Our estimated effects of returns to criminal and education experience are precisely estimated, but not particularly large in magnitude. This implies that the observed persistence in choices does not come primarily through this channel, but via state dependence and individual heterogeneity instead. This has important policy implications as well. If returns to criminal experience were high, then individuals who had accumulated a lot of experience might be very difficult to deter from committing crimes in the future. But since we find these returns to be low, this suggests that there does not come a point at which it is "too late" to intervene. Even youth who have amassed a long history of bad decisions can be affected by temporary interventions to break the state dependence and through changes to individual heterogeneity, such as reducing drug use or improving social/emotional skills.

Finally, it is important to stress that we are studying youth who have already committed somewhat serious criminal offenses. We feel that this is a particularly relevant group to study, as they represent a large proportion of overall youth crime, particularly serious crime. Furthermore, this is a group that has been studied relatively less intensively in the literature, largely due to data constraints. However, one implication of this is that our results do not necessarily generalize to the population at large. The factors that cause these serious offenders to reduce crime may not be the same as those that prevent people from committing their first crime. Additionally, what helps to reduce serious crimes such as robbery and assault, may not be as useful for preventing less serious crimes such as shoplifting.

Acknowledgments

This research was supported by the Social Sciences and Humanities Research Council of Canada. We would like to thank Shawn Bushway, two anonymous referees, and especially the editor for helpful comments and suggestions.

References

Agnew, Robert, Timothy Brezina, John Paul Wright, and Francis T. Cullen. 2002. "Strain, Personality Traits, and Delinquency: Extending General Strain Theory." *Criminology* 40 (1):43–72.

Anderson, D. Mark. 2014. "In School and Out of Trouble: The Minimum Dropout Age and Juvenile Crime." *Review of Economics and Statistics* 96 (2):318–331.

Bachman, Ronet, Raymond Paternoster, and Sally Ward. 1992. "The Rationality of Sexual Offending: Testing a Deterrence/Rational Choice Conception of Sexual Assault." *Law and Society Review* 26 (2):343–72.

Becker, Gary S. and Casey B. Mulligan. 1997. "The Endogenous Determination of Time Preference." *The Quarterly Journal of Economics* 112 (3):729–758.

Betts, Julian R. and Laurel L. McFarland. 1995. "Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments." *Journal of Human Resources* :741–765.

Blumstein, Alfred, David P. Farrington, and Soumyo Moitra. 1985. "Delinquency careers: Innocents, desisters, and persisters." In *Crime and Justice: An Annual Review of Research*, vol. 6, edited by M. Tonry and N. Morri. Chicago: University of Chicago Press, 187–2.

Brame, Robert, Shawn D Bushway, Raymond Paternoster, and Terence P Thornberry. 2005. "Temporal Linkages in Violent and Nonviolent Criminal Activity." *Journal of Quantitative Criminology* 21 (2):149–174.

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman. 2003. "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of

Uncertainty on College Choice." *International Economic Review* 44 (2):361–422. 2001 Lawrence R. Klein Lecture.

Caspi, Avshalom, Terrie E. Moffitt, Phil A. Silva, Magda Stouthamer-Loeber, Robert F. Krueger, and Pamela S. Schmutte. 1994. "Are Some People Crime Prone? Replication of the Person-Crime Relationship Across Countries, Genders, Race, and Methods." *Criminology* 32 (2):163–196.

Cawley, John, James J. Heckman, and Edward J. Vytlacil. 2001. "Three Observations on Wages and Measured Cognitive Ability." *Labour Economics* 8 (4):419–442.

Chen, M. Keith. 2013. "The Effect of Language on Economic Behavior: Evidence from Saving Rates, Health Behaviors and Retirement Access." *American Economic Review* 103 (2):690–731.

Chiteji, Ngina. 2010. "Time Preference, Noncognitive Skills and Well Being Across the Life Course: Do Noncognitive Skills Encourage Healthy Behavior?" *American Economic Review* 100 (2):200.

Cobb-Clark, Deborah A. and Michelle Tan. 2011. "Noncognitive Skills, Occupational Attainment, and Relative Wages." *Labour Economics* 18 (1):1–13.

Cooley, Jane, Salvador Navarro, and Yuya Takahashi. 2015. "Identification and Estimation of Time-varying Treatment Effects: How The Timing of Grade Retention Affects Outcomes." *Journal of Labor Economics* Forthcoming.

Cunha, Flavio, James J. Heckman, Lance J. Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, vol. 1, edited by Eric A. Hanushek and Frank Welch, chap. 12. Elsevier, 697–812.

Davis, Michael L. 1988. "Time and Punishment: An Intertemporal Model of Crime." *Journal of Political Economy* 96 (2):383–390.

Dellas, Harris and Plutarchos Sakellaris. 2003. "On the Cyclicality of Schooling: Theory and Evidence." *Oxford Economic Papers* 55 (1):148–172.

Dornbusch, Sanford M., J. Merrill Carlsmith, Steven J. Bushwall, Philip L. Ritter, Herbert Leiderman, Albert H. Hastorf, and Ruth T. Gross. 1985. "Single Parents, Extended Households, and the Control of Adolescents." *Child Development* 56 (2):326–341.

Durlauf, Steven N., Salvador Navarro, and David A. Rivers. 2008. "On the Interpretation of Aggregate Crime Regressions." In *Understanding Crime Trends*. Washington DC: National Academies Press.

———. 2010. "Understanding Aggregate Crime Regressions." *Journal of Econometrics* 158 (2):306–317.

Farrington, David P. 1986. "Age and Crime." In *Crime and Justice: An Annual Review of Research*, vol. 7, edited by M. Tonry and N. Morris. University of Chicago Press.

Farrington, David P., Bernard Gallagher, Lynda Morley, Raymond J. St Ledger, and Donald J. West. 1986. "Unemployment, School Leaving and Crime." *British Journal of Criminology* 26 (4):335–356.

Freeman, Richard B. 1996. "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?" *Journal of Economic Perspectives* 10 (1):25–42.

Geller, Amanda, Irwin Garfinkel, Carey E. Cooper, and Ronald B. Mincy. 2009. "Parental Incarceration and Child Well-being: Implications for Human Beings." *Social Science Quarterly* 90 (5):1186–1202.

Gottfredson, Michael R. and Travis Hirschi. 1990. *A General Theory of Crime*. Stanford University Press.

Gould, Eric D., Bruce A. Weinberg, and David B. Mustard. 2002. "Crime Rates and Local Labor Market Opportunities in the United States: 1979–1997." *Review of Economics and Statistics* 84 (1):45–61.

Grogger, Jeffrey. 1995. "The Effect of Arrests on the Employment and Earnings of Young Men." *The Quarterly Journal of Economics* 110:51–71.

———. 1998. "Market Wages and Youth Crime." *Journal of Labor Economics* 16 (4):756–791.

Hansen, David. 2011. *The Effects of Crime on Educational Investment: A Policy Simulation Approach*. Ph.D. thesis, Stanford University.

Heckman, James J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46 (4):931–959.

———. 1981. "Heterogeneity and State Dependence." In *Studies in Labor Markets*, edited by Sherwin Rosen, National Bureau of Economic Research. University of Chicago Press, 91–139.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3):411–482.

Herrnstein, Richard J. and Charles A. Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

Hill, Patrick L., Brent W. Roberts, Jeffrey T. Grogger, Jonathan Guryan, and Karen Sixkiller. 2011. "Decreasing Delinquency, Criminal Behavior, and Recidivism by Intervening on Psychological Factors Other than Cognitive Ability: A Review of the Interventional Literature." Tech. rep., NBER.

Hirschi, Travis and Michael Gottfredson. 1983. "Age and the Explanation of Crime." *American Journal of Sociology* 89:552–584.

Hjalmarsson, Randi, Helena Holmlund, and Matthew J. Lindquist. 2015. "The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Micro-Data." *The Economic Journal* 125 (587):1290–1326.

Imai, Susumu, Hajime Katayama, and Kala Krishna. 2006. "Crime and Young Men: The Role of Arrest, Criminal Experience and Heterogeneity." Tech. Rep. 12221, NBER.

Imai, Susumu and Kala Krishna. 2004. "Employment, Deterrence and Crime in a Dynamic Model." *International Economic Review* 45 (3):845–872.

Jacob, Brian A. and Lars Lefgren. 2003. "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime." *American Economic Review* 93 (5):1560–1577.

Kim, Dongwoo. 2014. *The Effect of Violent Crime on Educational Attainment*. Ph.D. thesis, University of Illinois at Urbana-Champaign.

Kling, Jeffrey R. 2006. "Incarceration Length, Employment, and Earnings." *American Economic Review* 96 (3):863–876.

Kotlarski, Ignacy I. 1967. "On Characterizing the Gamma and Normal Distribution." *Pacific Journal of Mathematics* 20:69–76.

Lewbel, Arthur. 2007. "Coherence and Completeness of Structural Models Containing a Dummy Endogenous Variable." *International Economic Review* 48 (4):1379–1392.

Lochner, Lance J. 2004. "Education, Work, and Crime: A Human Capital Approach." *International Economic Review* 45 (3):811–843.

———. 2007. "Individual Perceptions of the Criminal Justice System." *American Economic Review* 97 (1):444–460.

———. 2010. "Education and Crime." In *International Encyclopedia of Education*, edited by B. McGraw, P. Peterson, and E. Baker. Amsterdam: Elsevier, 3rd. ed.

———. 2011a. "Education Policy and Crime." In *Controlling Crime: Strategies and Tradeoffs.*, edited by P. Cook, J. Ludwig, and J. McCrary, chap. 11. University of Chicago Press.

———. 2011b. "Nonproduction Benefits of Education: Crime, Health, and Good Citizenship." In *Handbook of the Economics of Education*, vol. 4, chap. 2. Elsevier, 183–282.

Lochner, Lance J. and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* 94 (1):155–189.

Loughran, Thomas A., Holly Nguyen, Alex R. Piquero, and Jeffrey Fagan. 2013. "The Returns to Criminal Capital." *American Sociological Review* 78 (6):925–948.

Luallen, Jeremy. 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, Stephen, Olivier Marie, and Sunčica Vujić. 2011. "The Crime Reducing Effect of Education." *The Economic Journal* 121 (552):463–484.

Mastrobuoni, Giovanni and David A. Rivers. 2015. "Criminal Discount Factors and Deterrence." Working paper, University of Western Ontario, Department of Economics.

Merlo, Antonio and Kenneth I. Wolpin. 2015. "The Transition From School to Jail: Youth Crime and High School Completion Among Black Males." *European Economic Review* 79:234–251.

Mocan, H. Naci and Daniel I. Rees. 2005. "Economic Conditions, Deterrence and Juvenile Crime: Evidence from Micro Data." *American Law and Economics Review* 7 (2):319–349.

Monahan, Kathryn C., Laurence Steinberg, Elizabeth Cauffman, and Edward P. Mulvey. 2009. "Trajectories of Antisocial Behavior and Psychosocial Maturity from Adolescence to Young Adulthood." *Developmental Psychology* 45 (6):1654–1668.

Mulvey, Edward P. and Carol A. Schubert. 2012. "Some Initial Policy Implications from the Pathways to Desistance Study." *Victims and Offenses* 7 (4):407–427.

Munyo, Ignacio. 2015. "The Juvenile Crime Dilemma." *Review of Economic Dynamics* 18 (2):201–211.

Murnane, Richard J., John B. Willett, and Frank Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77 (2):251–266.

Nagin, Daniel, David P. Farrington, and Terrie E. Moffitt. 1995. "Life-Course Trajectories of Different Types of Offenders." *Criminology* 33 (1):111–139.

Nagin, Daniel and Kenneth C. Land. 1993. "Age, Criminal Careers, and Population Heterogeneity: Specification and Estimation of a Nonparametric, Mixed Poisson Model." *Criminology* 31 (3):327–362.

Nagin, Daniel and Raymond Paternoster. 1991. "On the Relationship of Past to Future Participation in Delinquency." *Criminology* 29 (2):163–189.

———. 2000. "Population Heterogeneity and State Dependence: State of the Evidence and Directions for Future Research." *Journal of Quantitative Criminology* 16 (2):117–144.

Nagin, Daniel and Joel Waldfogel. 1995. "The Effects of Criminality and Conviction on the Labor Market Status of Young British Offenders." *International Review of Law and Economics* 15 (1):109–126.

O'Donoghue, Ted and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review* 89 (1):103–124.

Paternoster, Raymond and Sally Simpson. 1996. "Sanction Threats and Appeals to Morality: Testing a Rational Choice Model of Corporate Crime." *Law and Society Review* 30 (3):549–83.

Piquero, Alex R, Carol A Schubert, and Robert Brame. 2014. "Comparing Official and Self-report RecordsofOffending across Gender and Race/Ethnicity in a Longitudinal Study of Serious Youthful Offenders." *Journal of Research in Crime and Delinquency* 51:526–555.

Pogarsky, Greg and Alex R. Piquero. 2003. "Can Punishment Encourage Offending?" *Crime and Delinquency* 40 (1):95–120.

Raphael, Steven and Rudolf Winter-Ebmer. 2001. "Identifying the Effect of Unemployment on Crime." *Journal of Law and Economics* 44 (1):259–83.

Sampson, Robert J. and John H. Laub. 2003. "Life-Course Desisters? Trajectories of Crime Among Delinquent Boys Followed to Age 70." *Criminology* 41 (3):555–592.

Schubert, Carol A., Edward P. Mulvey, Laurence Steinberg, Elizabeth Cauffman, Sandra H. Losoya, Thomas Hecker, Laurie Chassin, and George P. Knight. 2004. "Operational Lessons from the Pathways to Desistance Project." *Youth Violence and Juvenile Justice* 2 (3):237–255.

Tauchen, Helen, Ann Dryden Witte, and Harriet Griesinger. 1994. "Criminal Deterrence: Revisiting the Issue with a Birth Cohort." *Review of Economics and Statistics* 76 (3):399–412.

Usher, Dan. 1997. "Education as a Deterrent to Crime." *Canadian Journal of Economics* 30 (2):367–384.

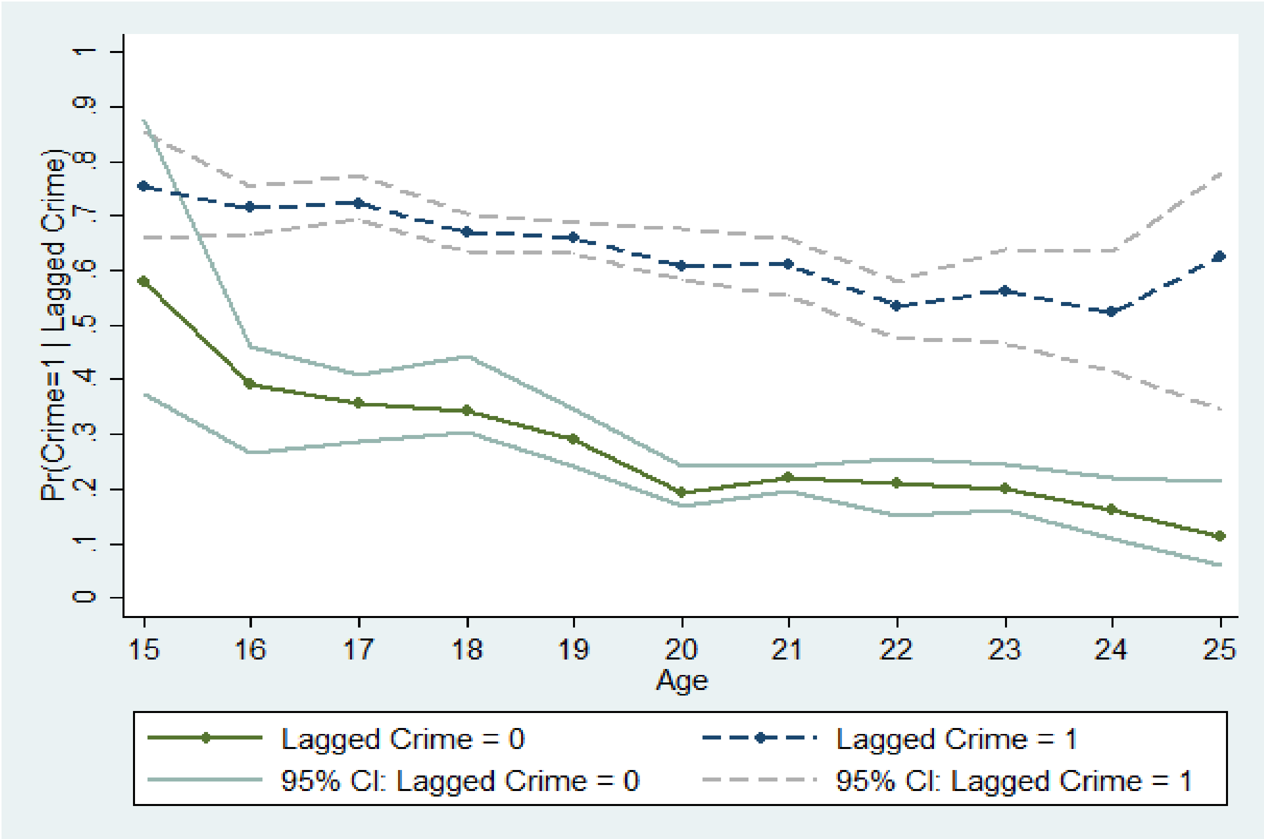
Waldfogel, Joel. 1994. "Does Conviction Have a Persistent Effect on Income and Employment?" *International Review of Law and Economics* 14 (1):103–119.

Wilson, James Q. and Richard J. Herrnstein. 1985. *Crime and Human Nature*. Simon and Schuster.

Witte, Ann Dryden and Helen Tauchen. 1994. "Work and Crime: An Exploration Using Panel Data." *Public Finance* 49:1–14.

Zelekha, Yaron and Ohad Bar-Efrat. 2011. "Crime, Terror and Corruption and Their Effect on Private Investment in Israel." *Open Economies Review* 22 (4):635–648.

Figure 1: Probability of Crime by Lagged Crime Choice and Age

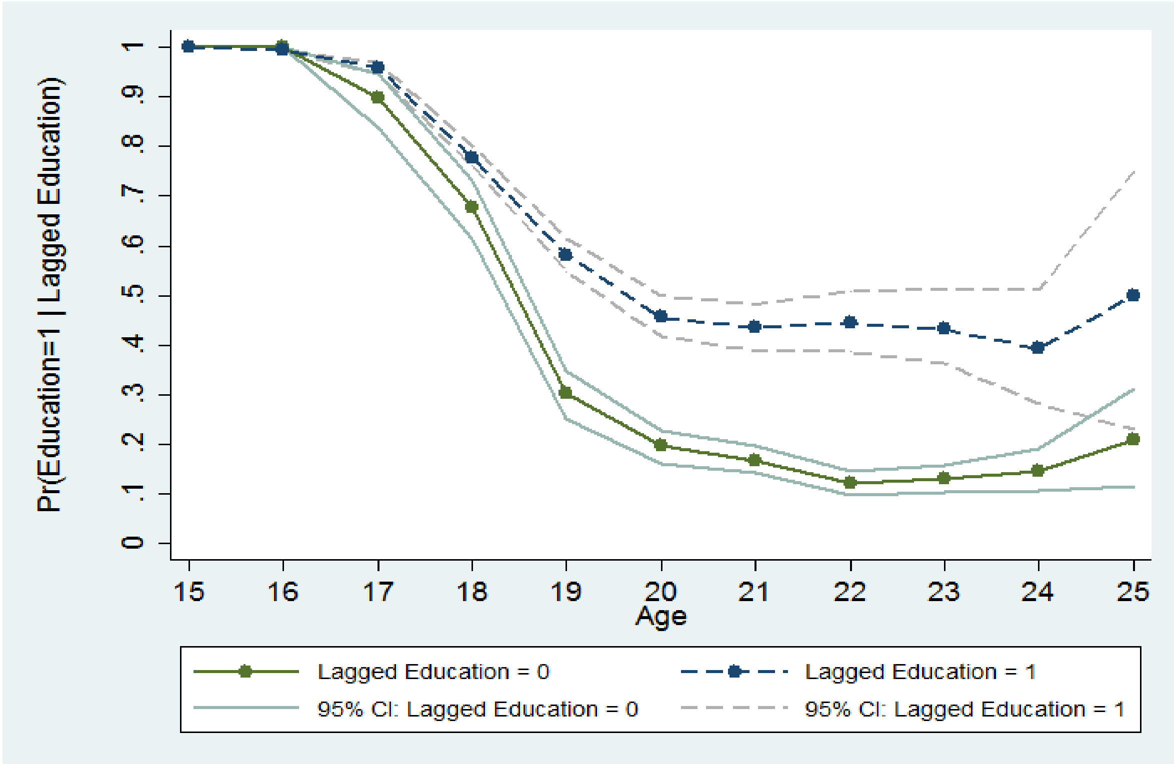


Notes:

1. The figures are based on the overall crime category. For each age category we run a probit of crime on lagged crime. We then predict the probability of engaging in crime by lagged crime and age. The confidence intervals are generated via bootstrapping.

2. Individuals can be considered for the study if they are found guilty of a misdemeanor weapons crime, which we do not categorize into one of our crime types (violent, property, drug). As a consequence, there are a small number of individuals with lagged crime equal to zero at age 15, even though all 15-year-olds entered the survey in the previous year.

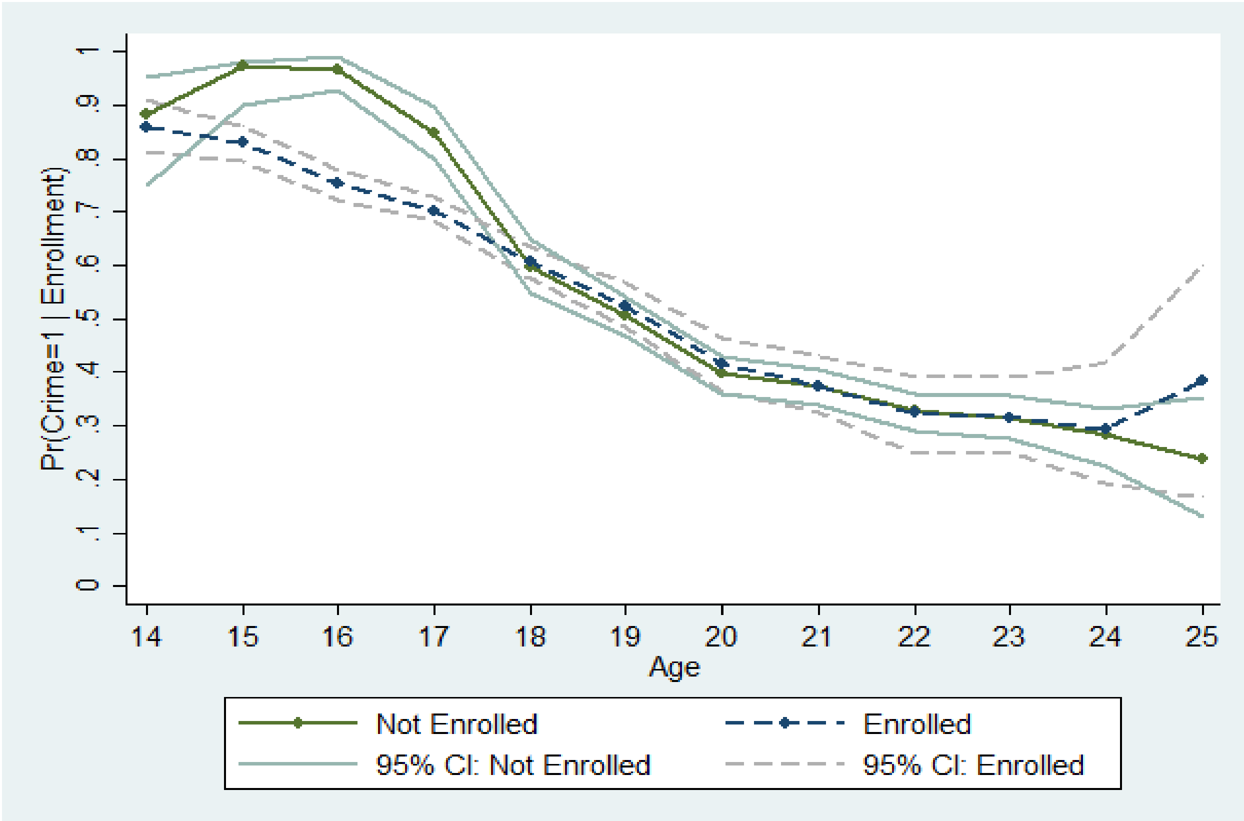
Figure 2: Probability of Education by Lagged Education Choice and Age



Note:

1. The figures are based on the overall crime category. For each age category we run a probit of education on lagged education. We then predict the probability of education by lagged education and age. The confidence intervals are generated via bootstrapping.

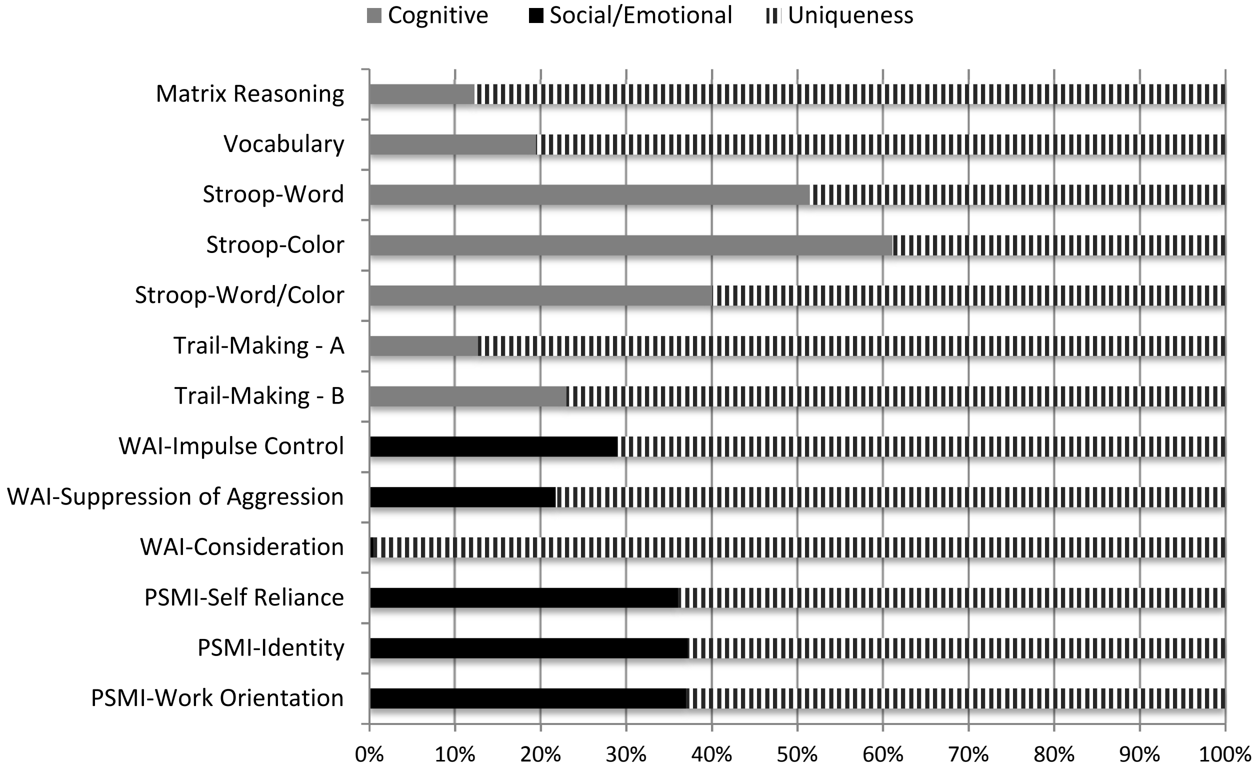
Figure 3: Probability of Crime by Enrollment Status and Age



Note:

1. The figures are based on the overall crime category. For each age category we run a probit of crime on enrollment. We then predict the probability of engaging in crime by enrollment and age. The confidence intervals are generated via bootstrapping.

Figure 4: Fraction of the Variance Explained by Skills

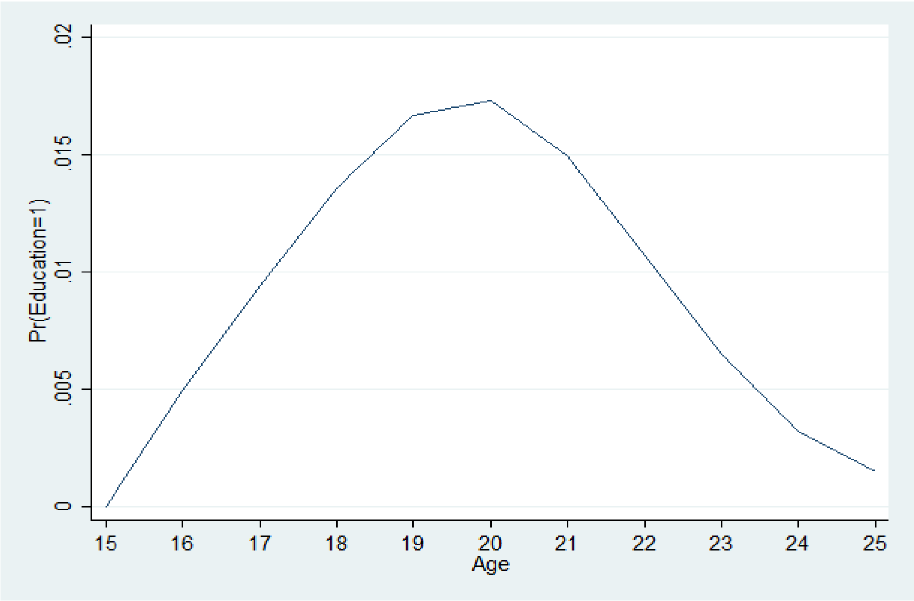


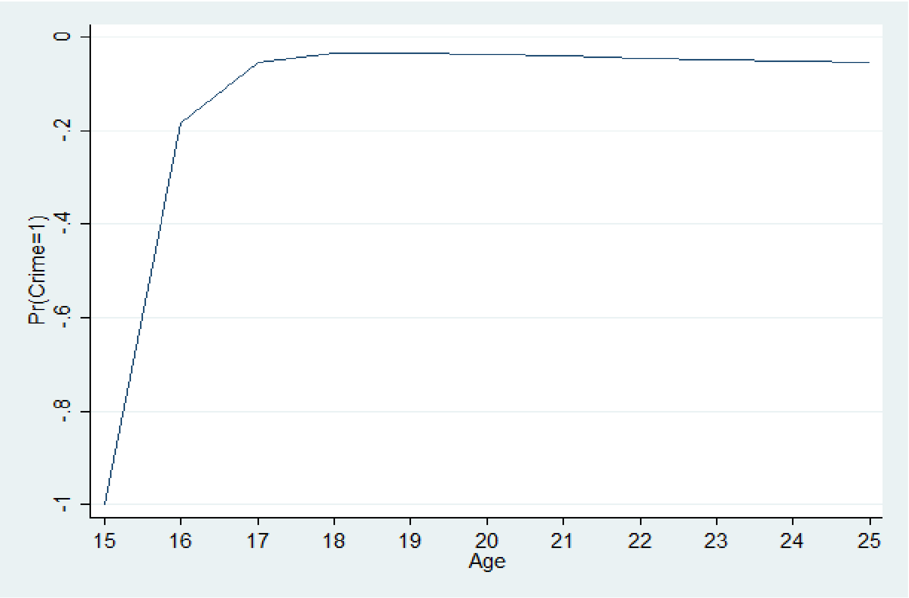
Notes:

1. We estimate a two factor model with cognitive and social/emotional measures. For the cognitive system, the components of WASI and Stroop are modeled using a linear in parameters specification of the form: $M_{j,i,t_i}^{cog} = x_{i,t_i}\beta_j^{cog} + \theta_i^{cog}\delta_{j,t_i}^{cog} + \xi_{j,i,t_i}^{cog}$, where j indexes the measure and i the individual. For the case of Trail-Making we use an ordered model of the form: $M_{j,i,t_i}^{cog} = \ell \Rightarrow \mathbb{1}\left(\psi_{j,\ell-1} < x_{i,t_i}\beta_j^{cog} + \theta_i^{cog}\delta_{j,t_i}^{cog} + \xi_{j,i,t_i}^{cog} \leq \psi_{j,\ell}\right)$.

2. For the social/emotional measures we use a linear in parameters specification of the form: $M_{k,i,t}^{emo} = x_{i,t}\beta_{k,t}^{emo} + \theta_i^{emo}\delta_{k,t}^{emo} + \xi_{k,i,t}^{emo}$, where k indexes the measure, i the individual and t age. The figure presents the average fraction of the variance explained by skills. For example, the fraction of the variance of test j explained by cognitive skills is given by: $\frac{1}{T_{cog}}\sum_t \frac{\sigma_{\theta,cog}^2(\delta_{j,t}^{cog})^2}{\sigma_{\theta,cog}^2(\delta_{j,t}^{cog})^2 + \sigma_{\xi,cog,j}^2}$, where T_{cog} is the number of ages for which we observe cognitive scores.

Figure 5: No Crime at Age 15 - Effect on Average Probability of Education and Crime





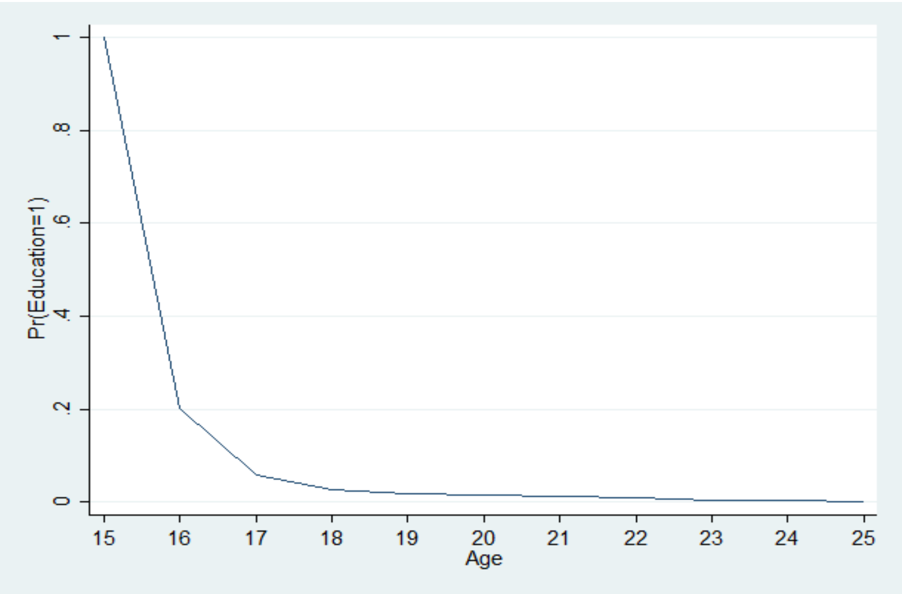
Notes:

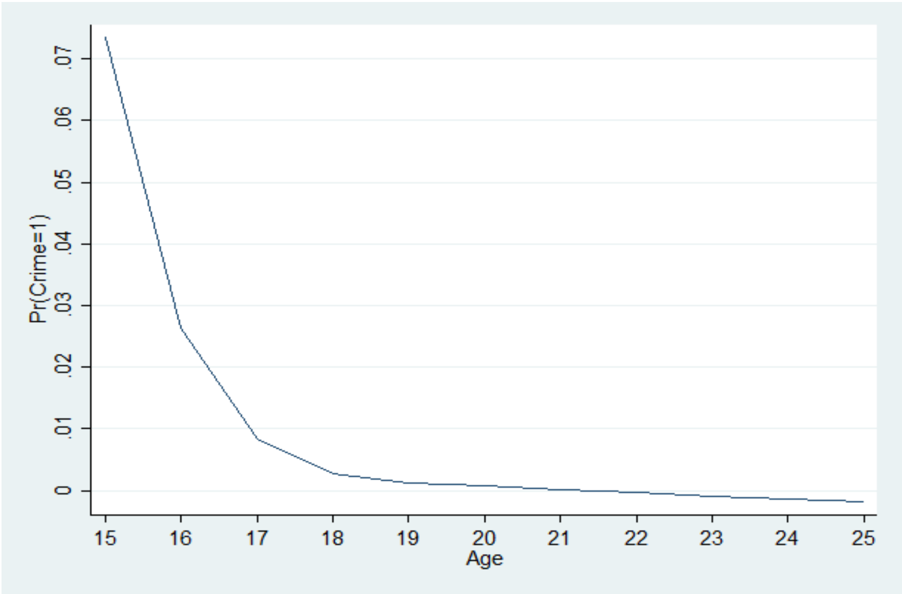
1. The figures are based on the overall crime category.

2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

3. Note that for the second figure, the comparison between two identical individuals who differ only along one dimension (crime) at age 15 implies that the average difference in the probability of crime between them is equal to -1 at that age by construction.

Figure 6: Enrolled at Age 15 - Effect on Average Probability of Education and Crime

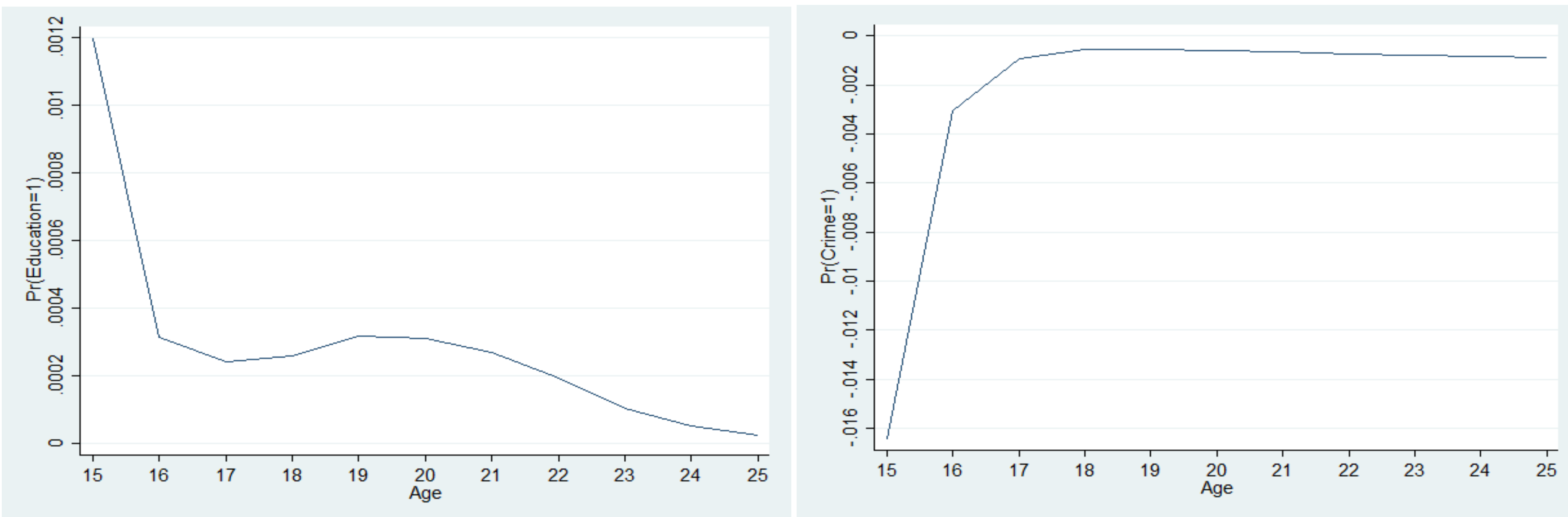




Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.
3. Note that for the first figure, the comparison between two identical individuals who differ only along one dimension (enrollment) at age 15 implies that the average difference in the probability of enrollment between them is equal to -1 at that age by construction.

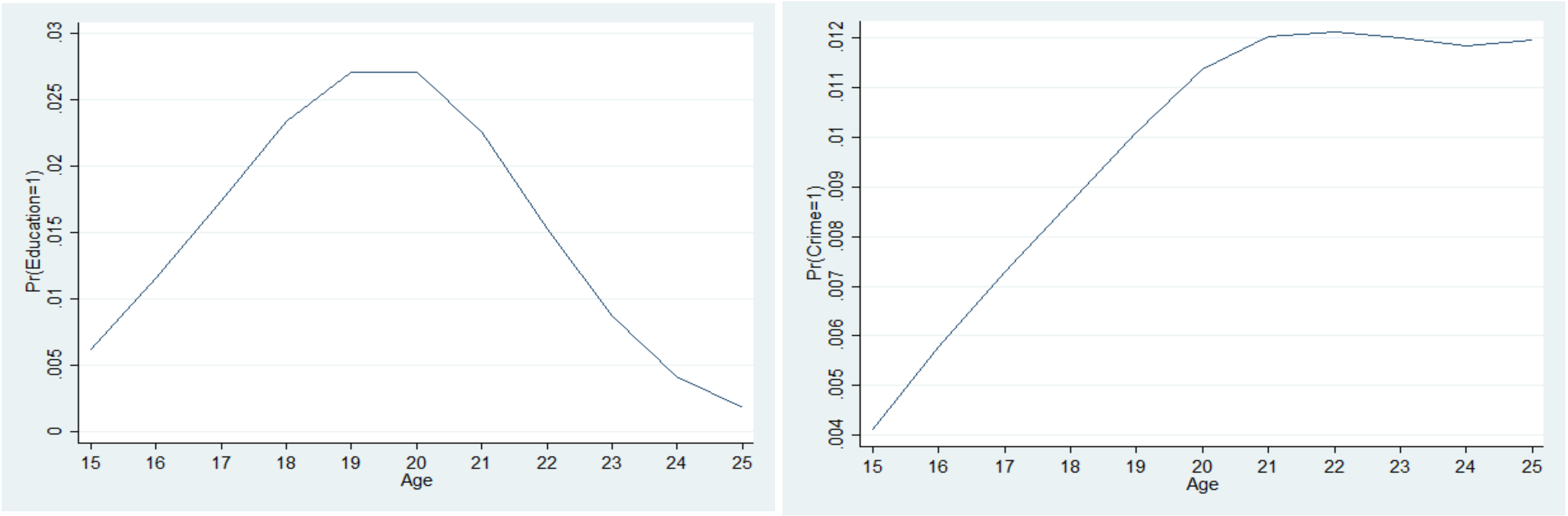
Figure 7: Increase in Certainty of Punishment at Age 15 - Effect on Average Probability of Education and Crime



Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

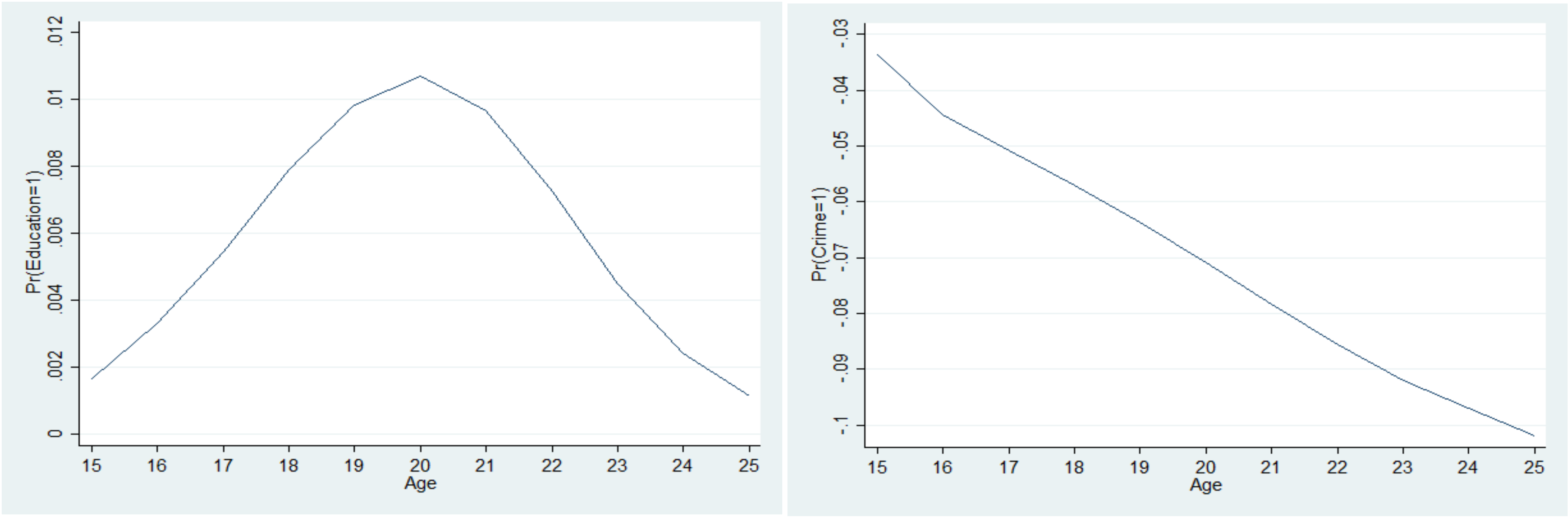
Figure 8: Cognitive Factor 25th versus 75th Percentile - Effect on Average Probability of Education and Crime



Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

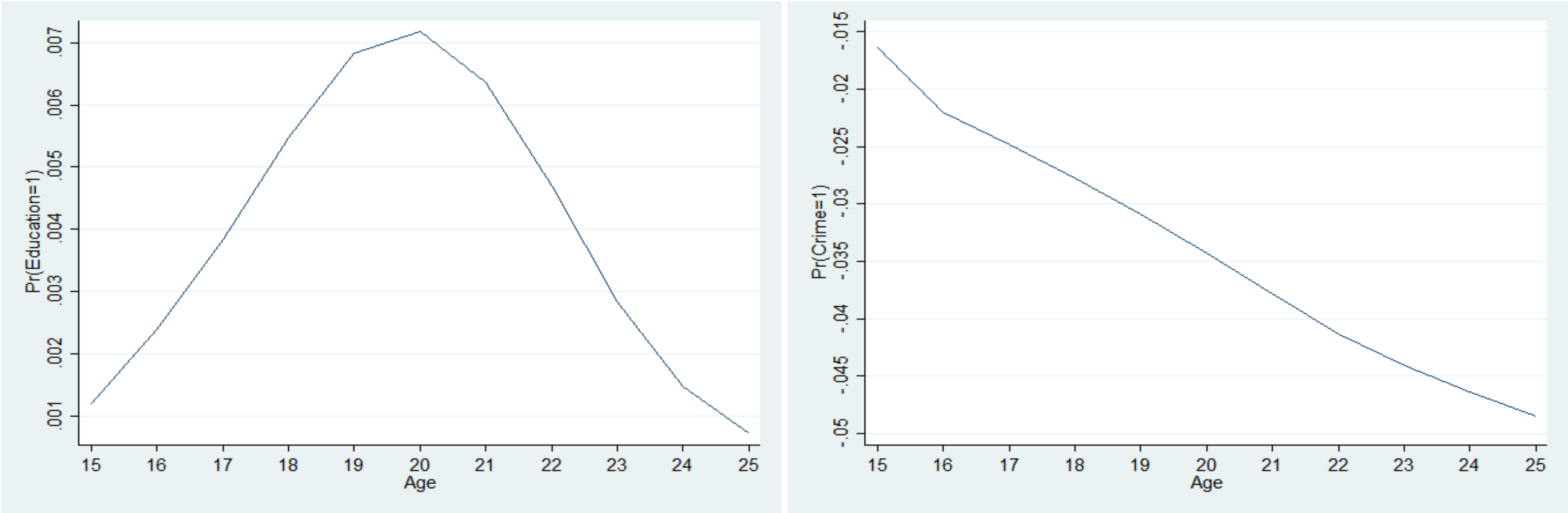
Figure 9: Social/Emotional Factor 25th versus 75th Percentile - Effect on Average Probability of Education and Crime



Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

Figure 10: Increase in Certainty of Punishment (Permanent) - Effect on Average Probability of Education and Crime



Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

Table 1: Pathways to Desistance Descriptive Statistics - Mean and Standard Deviation By Sample

Variable	All Crime		Drug-Related Crime		Violent Crime		Property Crime	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Age First Crime*	10.43	*1.80*	13.89	*1.68*	10.75	*2.00*	11.51	*2.21*
Age First Interview*	16.03	*1.14*	16.03	*1.14*	16.03	*1.14*	16.03	*1.14*
Phoenix*	0.49	*0.50*	0.50	*0.50*	0.49	*0.50*	0.49	*0.50*
Hispanic*	0.34	*0.47*	0.34	*0.47*	0.34	*0.47*	0.34	*0.47*
Black*	0.40	*0.49*	0.40	*0.49*	0.40	*0.49*	0.40	*0.49*
Other*	0.05	*0.21*	0.05	*0.21*	0.05	*0.21*	0.05	*0.21*
Female*	0.14	*0.35*	0.14	*0.35*	0.14	*0.35*	0.14	*0.35*
Siblings*	4.09	*2.41*	4.08	*2.41*	4.09	*2.41*	4.09	*2.41*
Non-Intact Family*	0.85	*0.35*	0.85	*0.35*	0.85	*0.35*	0.85	*0.35*
Individuals*	1185		1168		1188		1191	
Children	0.44	*0.82*	0.44	*0.81*	0.45	*0.82*	0.45	*0.82*
Family Crime	0.19	*0.40*	0.20	*0.40*	0.19	*0.40*	0.19	*0.39*
Certainty of Punishment	5.58	*2.32*	5.59	*2.33*	5.58	*2.32*	5.58	*2.32*
Drug Use	0.47	*0.50*	0.47	*0.50*	0.47	*0.50*	0.47	*0.50*
Local Unemployment Rate (%)	5.80	*1.56*	5.78	*1.56*	5.81	*1.55*	5.82	*1.55*
Future Outlook Inventory	2.59	*0.54*	2.59	*0.54*	2.59	*0.54*	2.59	*0.54*
Crime Rate	0.54	*0.50*	0.21	*0.41*	0.44	*0.50*	0.29	*0.45*
Enrollment Rate	0.54	*0.50*	0.54	*0.50*	0.54	*0.50*	0.54	*0.50*
Years of Education at Age 19	11.49	*1.31*	11.48	*1.31*	11.50	*1.31*	11.49	*1.31*
Observations	7376		7210		7424		7422	

Notes:

* Indicates variables that do not vary over time. Summary statistics for these variables are calculated using only the baseline survey.

1. The descriptive statistics reported in this table correspond to data from the combined baseline and follow-up surveys. Each observation is an individual-year pair.
2. The number of observations varies across the four samples since they differ in the number of missing values for each self-reported crime.
3. The crime and enrollment rates reflect the fraction of observations engaged in crime and enrolled in school, respectively.

Table 2: Pathways to Desistance Descriptive Statistics: Measures of Cognitive Skills

IQ and Components

Percentile	Score		
	IQ	**Vocabulary**	**Reasoning**
1%	55	20	20
5%	62	20	20
10%	67	24	23
25%	76	30	35
50%	85	38	44
75%	94	43	51
90%	102	51	55
95%	106	53	57
99%	115	61	61

Trail-Making

	% Sample
Part A	
Perfectly Normal	41.36
Normal	37.74
Mild / Moderately Impaired	13.56
Moderately / Severely Impaired	7.33
Part B	
Perfectly Normal	34.63
Normal	27.38
Mild / Moderately Impaired	26.37
Moderately / Severely Impaired	11.63

Stroop

	% Score < 40
Color	52.06
Word	36.31
Color/Word	20.89

Notes:

1. The descriptive statistics are based on the overall crime sample.
2. The estimate of general intellectual ability (IQ) is based on two subsets: Vocabulary and Matrix Reasoning.
3. The Trail-Making test is a measure of general brain function. Part A involves a series of numbers and the participant is required to connect the numbers in sequential order; Part B involves a series of numbers and letters and the participant is required to alternately connect letters and numbers in sequential order. The scores take one of four values, where the lowest two values indicate either mild/moderate impairment or moderate/severe impairment.
4. The Stroop Color/Word Test is used to examine the effects of interference on reading ability. The test has three parts, which relate to interference from words, colors, and both words and colors. The tests take a continuum of values, and for each test scores above 40 are considered “normal”.

Table 3: Estimated Parameters from Factor Analysis - Cognitive Skills

		WASI		Stroop			Trail-Making	
		Matrix	**Vocabulary**	**Word**	**Color**	**Color/Word**	**A**	**B**
Constant	**Age 14**	0.372*	-0.166	-0.308	-0.175	-0.138	0.000	0.000
		(0.222)	*(0.206)*	*(0.208)*	*(0.221)*	*(0.218)*	-	-
	Age 15	0.245	-0.157	-0.115	0.048	0.102	-0.644***	-0.379**
		(0.198)	*(0.187)*	*(0.188)*	*(0.193)*	*(0.194)*	*(0.155)*	*(0.171)*
	Age 16	0.309	-0.070	-0.046	0.199	0.127	-0.910***	-0.554***
		(0.199)	*(0.185)*	*(0.189)*	*(0.200)*	*(0.188)*	*(0.141)*	*(0.152)*
	Age 17	0.512**	-0.157	-0.069	0.248	0.276	-0.934***	-0.677***
		(0.206)	*(0.182)*	*(0.195)*	*(0.201)*	*(0.197)*	*(0.143)*	*(0.152)*
	Age 18	0.546**	0.025	0.077	0.397	0.242	-0.858***	-0.776***
		(0.236)	*(0.215)*	*(0.246)*	*(0.256)*	*(0.224)*	*(0.219)*	*(0.202)*
Phoenix		0.332***	0.743***	0.346***	0.141	0.269***	-0.387***	-0.363***
		(0.093)	*(0.086)*	*(0.096)*	*(0.096)*	*(0.090)*	*(0.116)*	*(0.125)*
Hispanic		-0.412***	-0.652***	-0.220**	-0.254***	-0.264***	0.326***	0.436***
		(0.104)	*(0.092)*	*(0.096)*	*(0.094)*	*(0.092)*	*(0.122)*	*(0.131)*
Black		-0.465***	-0.328***	-0.250**	-0.175	-0.335***	0.467***	0.429***
		(0.108)	*(0.103)*	*(0.111)*	*(0.113)*	*(0.104)*	*(0.141)*	*(0.152)*
Other		-0.268	-0.456**	-0.239	-0.378**	-0.422**	0.189	0.256
		(0.187)	*(0.180)*	*(0.184)*	*(0.184)*	*(0.185)*	*(0.232)*	*(0.254)*
Female		-0.023	0.004	0.179*	0.090	0.048	-0.101	-0.221
		(0.110)	*(0.096)*	*(0.096)*	*(0.098)*	*(0.095)*	*(0.124)*	*(0.135)*
Siblings		-0.015	-0.023*	-0.016	-0.016	-0.023	-0.024	-0.003
		(0.015)	*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.017)*	*(0.019)*
FOI		-0.068	0.080	0.055	-0.007	0.015	0.018	0.140*
		(0.066)	*(0.059)*	*(0.061)*	*(0.062)*	*(0.062)*	*(0.078)*	*(0.082)*
Cognitive Ability	**Age 14**	1.000	1.191	2.358**	2.594**	2.048**	-0.328	-0.430
		-	*(0.775)*	*(0.972)*	*(1.146)*	*(0.924)*	*(0.484)*	*(0.610)*
	Age 15	1.464*	1.641**	2.048**	2.150**	1.676**	-1.832*	-2.548**
		(0.752)	*(0.792)*	*(0.928)*	*(0.981)*	*(0.804)*	*(0.950)*	*(1.227)*
	Age 16	0.862*	1.367**	2.697**	3.078**	2.313**	-1.465**	-2.371**
		(0.444)	*(0.660)*	*(1.213)*	*(1.378)*	*(1.043)*	*(0.721)*	*(1.094)*
	Age 17	1.472**	1.199**	2.385**	2.769**	2.596**	-1.939**	-2.440**
		(0.686)	*(0.574)*	*(1.073)*	*(1.240)*	*(1.167)*	*(0.909)*	*(1.134)*
	Age 18	1.236	1.809*	3.552**	3.502**	2.307**	-0.909	-1.783*
		(0.790)	*(0.993)*	*(1.663)*	*(1.648)*	*(1.133)*	*(0.819)*	*(0.980)*
Variance		0.809***	0.659***	0.469***	0.371***	0.539***	1.000	1.000
		(0.048)	*(0.035)*	*(0.031)*	*(0.025)*	*(0.025)*	-	-
Cutoff 1		-	-	-	-	-	-0.964***	-0.550*
							(0.255)	*(0.287)*
Cutoff 2		-	-	-	-	-	0.238	0.304
							(0.256)	*(0.288)*
Cutoff 3		-	-	-	-	-	1.007***	1.400***
							(0.252)	*(0.292)*

Notes:

1. We estimate a two factor model with cognitive and social/emotional measures. The table presents the parameter estimates for the cognitive measure system. The components of WASI and Stroop are modeled using a linear in parameters specification of the form: $M^{cog}_{j,i,t_i} = x_{i,t_i}\beta^{cog}_j + \theta^{cog}_i \delta^{cog}_{j,t_i} + \xi^{cog}_{j,i,t_i}$, where *j* indexes the measure (column in the table) and *i* the individual. For the case of Trail-Making we use an ordered model of the form: $M^{cog}_{j,i,t_i} = \ell \Rightarrow \mathbb{1}\left(\psi_{j,\ell-1} < x_{i,t_i}\beta^{cog}_j + \theta^{cog}_i \delta^{cog}_{j,t_i} + \xi^{cog}_{j,i,t_i} \leq \psi_{j,\ell}\right)$.
2. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.

Table 4: Estimated Parameters from Factor Analysis - Social/Emotional Skills

		WAI			PSMI		
		Impulse Control	**Suppression of Aggression**	**Consideration of Others**	**Self Reliance**	**Identity**	**Work Orientation**
Constant	**Age 14**	-1.221***	-0.837***	-1.790***	-0.704***	-0.552***	-1.368***
		(0.115)	*(0.115)*	*(0.096)*	*(0.116)*	*(0.126)*	*(0.100)*
	Age 15	-1.093***	-0.749***	-1.960***	-0.673***	-0.722***	-1.291***
		(0.090)	*(0.090)*	*(0.076)*	*(0.102)*	*(0.103)*	*(0.090)*
	Age 16	-1.032***	-0.755***	-2.019***	-0.501***	-0.611***	-1.176***
		(0.079)	*(0.080)*	*(0.059)*	*(0.099)*	*(0.095)*	*(0.080)*
	Age 17	-1.000***	-0.751***	-1.968***	-0.390***	-0.540***	-1.121***
		(0.077)	*(0.081)*	*(0.063)*	*(0.095)*	*(0.096)*	*(0.079)*
	Age 18	-0.920***	-0.676***	-1.950***	-0.302***	-0.437***	-0.934***
		(0.082)	*(0.083)*	*(0.062)*	*(0.101)*	*(0.098)*	*(0.085)*
	Age 19	-0.880***	-0.605***	-1.891***	-0.211**	-0.382***	-0.839***
		(0.087)	*(0.084)*	*(0.067)*	*(0.105)*	*(0.098)*	*(0.084)*
	Age 20	-0.818***	-0.555***	-1.863***	-0.091	-0.297***	-0.708***
		(0.085)	*(0.086)*	*(0.068)*	*(0.108)*	*(0.105)*	*(0.086)*
	Age 21	-0.800***	-0.505***	-1.830***	-0.097	-0.299***	-0.701***
		(0.085)	*(0.086)*	*(0.066)*	*(0.103)*	*(0.103)*	*(0.089)*
	Age 22	-0.731***	-0.411***	-1.789***	-0.035	-0.247**	-0.634***
		(0.085)	*(0.089)*	*(0.068)*	*(0.108)*	*(0.104)*	*(0.086)*
	Age 23	-0.689***	-0.375***	-1.807***	0.017	-0.185*	-0.616***
		(0.091)	*(0.093)*	*(0.076)*	*(0.108)*	*(0.106)*	*(0.091)*
	Age 24	-0.674***	-0.377***	-1.837***	0.017	-0.169	-0.583***
		(0.102)	*(0.106)*	*(0.082)*	*(0.130)*	*(0.138)*	*(0.105)*
	Age 25	-0.582***	-0.438**	-1.688***	-0.159	-0.447*	-0.762***
		(0.203)	*(0.200)*	*(0.212)*	*(0.245)*	*(0.244)*	*(0.228)*
Phoenix		-0.222***	0.078**	-0.066***	-0.166***	-0.143***	-0.091**
		(0.041)	*(0.038)*	*(0.020)*	*(0.049)*	*(0.051)*	*(0.046)*
Hispanic		0.126***	-0.134***	-0.043**	-0.326***	-0.316***	-0.210***
		(0.041)	*(0.039)*	*(0.022)*	*(0.050)*	*(0.053)*	*(0.046)*
Black		0.327***	-0.128***	-0.062**	0.009	-0.036	-0.052
		(0.049)	*(0.047)*	*(0.024)*	*(0.059)*	*(0.062)*	*(0.055)*
Other		0.234***	-0.040	-0.012	-0.211**	-0.171*	-0.052
		(0.080)	*(0.077)*	*(0.045)*	*(0.096)*	*(0.100)*	*(0.055)*
Female		0.187***	0.135***	0.179***	0.141***	-0.029	-0.052
		(0.042)	*(0.037)*	*(0.022)*	*(0.050)*	*(0.053)*	*(0.055)*
Siblings		-0.006	0.006	0.008**	-0.002	-0.009	-0.052
		(0.006)	*(0.006)*	*(0.003)*	*(0.008)*	*(0.008)*	*(0.055)*
FOI		0.305***	0.229***	0.735***	0.156***	0.238***	-0.052
		(0.018)	*(0.018)*	*(0.017)*	*(0.024)*	*(0.023)*	*(0.055)*
Behavioral Ability	**Age 14**	1.000	0.924**	0.301	1.107***	1.148***	1.108***
		-	*(0.435)*	*(0.224)*	*(0.343)*	*(0.344)*	*(0.323)*
	Age 15	0.967***	0.880***	0.186	1.366***	1.309***	1.338***
		(0.313)	*(0.243)*	*(0.152)*	*(0.366)*	*(0.370)*	*(0.349)*
	Age 16	0.921***	0.854***	0.151	1.323***	1.388***	1.292***
		(0.263)	*(0.246)*	*(0.111)*	*(0.364)*	*(0.358)*	*(0.334)*
	Age 17	1.032***	0.948***	0.109	1.220***	1.257***	1.215***
		(0.277)	*(0.256)*	*(0.098)*	*(0.316)*	*(0.335)*	*(0.314)*
	Age 18	1.088***	0.997***	0.053	1.227***	1.181***	1.193***
		(0.292)	*(0.281)*	*(0.110)*	*(0.323)*	*(0.309)*	*(0.319)*
	Age 19	1.208***	1.095***	0.136	1.320***	1.362***	1.332***
		(0.328)	*(0.297)*	*(0.120)*	*(0.358)*	*(0.367)*	*(0.338)*
	Age 20	1.267***	1.148***	0.133	1.351***	1.429***	1.359***
		(0.344)	*(0.312)*	*(0.120)*	*(0.360)*	*(0.389)*	*(0.352)*
	Age 21	1.192***	1.110***	0.079	1.327***	1.413***	1.325***
		(0.316)	*(0.308)*	*(0.106)*	*(0.352)*	*(0.378)*	*(0.351)*
	Age 22	1.173***	1.109***	0.039	1.285***	1.350***	1.308***
		(0.308)	*(0.322)*	*(0.119)*	*(0.349)*	*(0.361)*	*(0.346)*
	Age 23	1.191***	1.008***	0.137	1.309***	1.276***	1.268***
		(0.329)	*(0.299)*	*(0.129)*	*(0.358)*	*(0.356)*	*(0.343)*
	Age 24	1.369***	1.229***	0.105	1.041***	1.068***	1.149***
		(0.388)	*(0.364)*	*(0.173)*	*(0.334)*	*(0.368)*	*(0.356)*
	Age 25	1.291**	0.954	0.088	1.792**	1.698*	1.293**
		(0.632)	*(0.607)*	*(0.545)*	*(0.804)*	*(0.945)*	*(0.609)*
Variance		0.609***	0.719***	0.805***	0.574***	0.562***	0.521***
		(0.010)	*(0.011)*	*(0.011)*	*(0.012)*	*(0.012)*	*(0.010)*

Notes:

1. We estimate a two factor model with cognitive and social/emotional measures. The table presents the parameter estimates for the social/emotional measure system. We use a linear in parameters specification of the form: $M^{emo}_{k,i,t} = x_{i,t}\beta^{emo}_{k,t} + \theta^{emo}_{i}\delta^{emo}_{k,t} + \xi^{emo}_{k,i,t}$, where k indexes the measure (column in the table), i the individual, and t age.
2. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.

Table 5: Average Marginal Effects from Probits for Crime and Education (Overall Crime)

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.049**	0.041**	0.038**	0.030*	0.050**	0.049**	0.051**	0.049**	0.096***	0.044**	0.039*	0.039*
	(0.021)	*(0.020)*	*(0.019)*	*(0.018)*	*(0.021)*	*(0.019)*	*(0.021)*	*(0.021)*	*(0.018)*	*(0.020)*	*(0.022)*	*(0.020)*
Hispanic	-0.025*	-0.020			-0.025*	-0.022	-0.026*	-0.032**	-0.026*	-0.021	-0.013	-0.024
	(0.015)	*(0.015)*			*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*
Black	0.024	-0.030*			0.024	-0.029	0.042**	-0.046**	0.024	-0.030	0.039**	-0.029
	(0.017)	*(0.018)*			*(0.017)*	*(0.018)*	*(0.018)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.018)*	*(0.019)*
Other	0.034	-0.025			0.034	-0.023	0.038	-0.036	0.036	-0.024	0.042	-0.015
	(0.027)	*(0.030)*			*(0.027)*	*(0.030)*	*(0.028)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*
Female	0.058***	-0.101***	0.054***	-0.087***	0.058***	-0.098***	0.070***	-0.168***	0.057***	-0.100***	0.053***	-0.096***
	(0.015)	*(0.016)*	*(0.014)*	*(0.017)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*
Non-intact Family	-0.050***	0.031*			-0.051***	0.028*	-0.052***	0.040**	-0.051***	0.030*	-0.049***	0.031**
	(0.015)	*(0.016)*			*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*
Siblings	-0.002	0.004			-0.002	0.003	-0.003	0.006**	-0.002	0.004	-0.002	0.004
	(0.002)	*(0.003)*			*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.002)*
Age	-0.080***	-0.028***	-0.083***	-0.035***	-0.080***	-0.036***	-0.104***	-0.017*	-0.082***	-0.031***	-0.079***	-0.026***
	(0.004)	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.004)*	*(0.002)*	*(0.009)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.007)*
Certainty of Punishment	0.003	-0.022***			0.003	-0.022***	0.005*	-0.028***	0.003	-0.022***	0.002	-0.018***
	(0.003)	*(0.003)*			*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.018**	0.008			-0.017**	0.007	-0.032***	0.013*	-0.017**	0.007	-0.017**	0.003
	(0.007)	*(0.007)*			*(0.007)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Family Crime	0.002	0.149***			0.002	0.150***	-0.002	0.175***	0.004	0.149***	0.001	0.146***
	(0.015)	*(0.015)*			*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*
Drug Use	-0.001	0.224***			-0.001	0.225***	-0.009	0.267***	-0.000	0.224***	-0.001	0.204***
	(0.012)	*(0.010)*			*(0.012)*	*(0.010)*	*(0.012)*	*(0.012)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.011)*
Unemployment Rate	0.021***	0.011**	0.021***	0.010*	0.021***	0.014***	0.023***	0.009	0.037***	0.012**	0.021***	0.010*
	(0.006)	*(0.005)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.005)*
Future Outlook Inventory	0.019*	-0.024**			0.019*	-0.023**	0.024**	-0.030**	0.017	-0.024**	0.022*	0.016
	(0.011)	*(0.011)*			*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.012)*	*(0.013)*
Years of Crime	-0.007***	0.020***	-0.007***	0.039***	-0.007***	0.020***			-0.007***	0.020***	-0.007**	0.017***
	(0.003)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.003)*	*(0.003)*			*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Years of Education	0.006	-0.003	0.011***	-0.014***	0.007	-0.001			0.006	-0.002	0.006	-0.005
	(0.004)	*(0.004)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.004)*			*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Cognitive Factor	0.036	0.014			0.038*	0.017	0.041*	0.030	0.036	0.015		
	(0.023)	*(0.024)*			*(0.023)*	*(0.024)*	*(0.024)*	*(0.024)*	*(0.023)*	*(0.024)*		
Social/Emotional Factor	0.007	-0.080***			0.006	-0.080***	0.019	-0.127***	0.007	-0.080***		
	(0.014)	*(0.014)*			*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*		
Schools per Young Person	0.322***		0.323***		0.313***		0.319***				0.311***	
	(0.071)		*(0.072)*		*(0.071)*		*(0.072)*				*(0.071)*	
Lagged Enrollment	0.189***		0.191***		0.189***				0.190***		0.185***	
	(0.012)		*(0.012)*		*(0.012)*				*(0.012)*		*(0.012)*	
Enrollment		0.088*		0.083		0.025*		0.202***		0.065		0.096**
		(0.049)		*(0.053)*		*(0.014)*		*(0.063)*		*(0.051)*		*(0.047)*
Lagged Crime		0.158***		0.235***		0.159***				0.159***		0.142***
		(0.012)		*(0.013)*		*(0.012)*				*(0.012)*		*(0.012)*
WASI Reasoning Score											-0.002	-0.005
											(0.006)	*(0.007)*
WASI Vocabulary Score											0.000	0.001
											(0.007)	*(0.007)*
Stroop: Color/Word											0.004	-0.010
											(0.007)	*(0.007)*
Stroop: Word											0.009	-0.012
											(0.007)	*(0.008)*
Stroop: Color											-0.003	0.014*
											(0.008)	*(0.008)*

Table 5: Average Marginal Effects from Probits for Crime and Education (Overall Crime)

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Trail-Making: Part B											-0.016**	-0.007
											(0.007)	*(0.007)*
Trail-Making: Part A											-0.003	-0.002
											(0.007)	*(0.007)*
WAI - Impulse Response											-0.008	-0.030***
											(0.007)	*(0.008)*
WAI - Suppression of Aggression											0.010	-0.044***
											(0.007)	*(0.007)*
WAI - Consideration of Others											0.002	-0.027***
											(0.006)	*(0.006)*
PSMI - Self Reliance											-0.014	0.023**
											(0.010)	*(0.011)*
PSMI - Identity											0.036***	-0.018*
											(0.010)	*(0.011)*
PSMI - Work Orientation											-0.023**	-0.011
											(0.009)	*(0.010)*
Rho	-0.142		-0.137				-0.377***		-0.091		-0.157	
	(0.106)		(0.102)				(0.144)		(0.109)		*-0.105*	
Observations	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. The errors in the enrollment and crime equations are allowed to be correlated in every specification, expect for specification (3). Rho denotes the correlation in errors.
3. Every specification includes an exclusion restriction that enters the education equation only (schools per young person) except for the specification in column (5).
4. In column (2) we exclude the cognitive and social/emotional factors and control for just a few variables (location, gender, age, and local unemployment rate). In column (3) the errors in the enrollment and crime equations are uncorrelated. The specification in column (4) does not account for any dynamics in the crime and education equations (years of experience and state dependence). The specification in column (5) does not include the exclusion restriction. In the last column we replace our factor estimates of cognitive and social/emotional skills with the measures used to infer them.

Table 6: The Effect of Educational Attainment on Crime
Alternative Specifications

Reduced Set of Controls	**Full Sample**	**Males Only**	**Males Only, Alternative Measure of Education**	**Males Only, Alternative Measure of Education, Age 18+**
	(1)	(2)	(3)	(4)
Years of Education	-0.020***	-0.025***	-0.028***	-0.028***
	(0.005)	*(0.005)*	*(0.005)*	*(0.005)*
Full Set of Controls	**Full Sample**	**Males Only**	**Males Only, Alternative Measure of Education**	**Males Only, Alternative Measure of Education, Age 18+**
	(1)	(2)	(3)	(4)
Years of Education	-0.003	-0.004	-0.011**	-0.010**
	(0.004)	*(0.005)*	*(0.004)*	*(0.005)*
Observations	5,190	4,277	4,277	3,574

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.

2. In the first set of results we include only a reduced set of controls (location, non-intact family, age, unemployment rate, and IQ). In the second set we include the full set of controls from our baseline specification, including lagged decisions, experience, and skills.

3. In columns (3) and (4) we use an alternative measure of years of education that does not include schooling obtained in jail.

Appendix A: Additional Results

Table A1: Average Marginal Effects from Probits for Crime and Education (Overall Crime) - Robustness Checks 1

Variable	Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)	
	(1)		(2)		(3)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.071***	0.047***	0.057***	0.039**	0.055***	0.037**
	(0.021)	*(0.018)*	*(0.019)*	*(0.019)*	*(0.019)*	*(0.019)*
Hispanic	-0.035**	-0.025*	-0.005	-0.033**	-0.005	-0.033**
	(0.014)	*(0.015)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*
Black	0.000	-0.033*	0.042**	-0.042**	0.039**	-0.043**
	(0.017)	*(0.017)*	*(0.016)*	*(0.017)*	*(0.016)*	*(0.017)*
Other	0.009	-0.020	0.032	-0.026	0.031	-0.027
	(0.027)	*(0.028)*	*(0.026)*	*(0.028)*	*(0.026)*	*(0.028)*
Female	0.118***	-0.096***	0.019	-0.070***	0.018	-0.067***
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*
Non-intact Family	-0.058***	0.023	-0.030**	0.015	-0.030**	0.014
	(0.015)	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*
Siblings	-0.002	0.003	-0.001	0.003	-0.001	0.003
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.070***	-0.034***	-0.074***	-0.035***	-0.074***	-0.035***
	(0.004)	*(0.004)*	*(0.004)*	*(0.007)*	*(0.004)*	*(0.007)*
Certainty of Punishment	0.006**	-0.019***	-0.001	-0.017***	-0.001	-0.017***
	(0.003)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Children	-0.016**	0.008	-0.015**	0.011*	-0.013**	0.012*
	(0.008)	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Family Crime	-0.033**	0.144***	0.009	0.131***	0.007	0.132***
	(0.014)	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*
Drug Use	0.044***	0.226***	-0.040***	0.233***	-0.041***	0.231***
	(0.011)	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*
Unemployment Rate	0.020***	0.013***	0.023***	0.011**	0.023***	0.011**
	(0.006)	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*
Future Outlook Inventory	0.019*	-0.029***	0.016	-0.027**	0.016	-0.027**
	(0.011)	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*
Years of Crime	-0.007***	0.021***	-0.005*	0.020***	-0.005*	0.022***
	(0.003)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.003)*	*(0.003)*
Years of Education	0.023***	-0.002	-0.007*	0.001	-0.002	-0.001
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Schools per Young Person	0.169**		0.316***		0.317***	
	(0.072)		*(0.067)*		*(0.067)*	
Lagged Enrollment	0.174***		0.198***		0.223***	
	(0.013)		*(0.011)*		*(0.013)*	

Table A1: Average Marginal Effects from Probits for Crime and Education (Overall Crime) - Robustness Checks 1

Variable	**Choices While in Jail (1)**		**Choices While in Jail (2)**		**Choices While in Jail (3)**	
	(1)		(2)		(3)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime
Enrollment		0.055***				
		(0.016)				
Lagged Crime		0.157***		0.149***		0.162***
		(0.012)		*(0.012)*		*(0.014)*
Cognitive Factor	0.033	0.006	0.002	0.017	-0.001	0.018
	(0.022)	*(0.022)*	*(0.021)*	*(0.022)*	*(0.021)*	*(0.022)*
Social/Emotional Factor	0.019	-0.076***	-0.012	-0.072***	-0.013	-0.071***
	(0.013)	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*
Jail			0.100***	0.119***	0.357***	0.163*
			(0.012)	*(0.013)*	*(0.078)*	*(0.092)*
Enrollment (alternative)				0.055		0.061
				(0.047)		*(0.046)*
Years of Crime * Jail					0.003	-0.007
					(0.005)	*(0.005)*
Years of Education * Jail					-0.020***	0.005
					(0.007)	*(0.007)*
Lagged Enrollment * Jail					-0.089***	
					(0.023)	
Lagged Crime * Jail						-0.052**
						(0.026)
Enrollment * Jail						-0.032
						(0.026)
Rho	-0.068**		-0.074		-0.067	
	(0.033)		*(0.100)*		*(0.010)*	
Observations	6,189	6,189	6,189	6,189	6,189	6,189

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. In column (1) enrollment is set to zero if an individual did not attend a community school. In column (2), we condition on whether the individual is interviewed in jail, and in column (3) we interact the jail dummy with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail.

Table A2: Average Marginal Effects from Probits for Crime and Education (Overall Crime) - Robustness Checks 2

Variable	Excluding Drug Use		Enrollment Based on Attendance	
	(1)		(2)	
	Educ.	Crime	Educ.	Crime
Phoenix	0.049**	0.038*	0.017	0.045**
	(0.021)	*(0.021)*	*(0.018)*	*(0.019)*
Hispanic	-0.025*	-0.025	-0.021*	-0.018
	(0.015)	*(0.016)*	*(0.012)*	*(0.016)*
Black	0.024	-0.041**	0.002	-0.027
	(0.017)	*(0.019)*	*(0.013)*	*(0.018)*
Other	0.034	-0.042	-0.022	-0.022
	(0.027)	*(0.031)*	*(0.022)*	*(0.030)*
Female	0.058***	-0.097***	0.008	-0.095***
	(0.015)	*(0.017)*	*(0.012)*	*(0.016)*
Non-intact Family	-0.050***	0.028*	-0.004	0.027*
	(0.015)	*(0.017)*	*(0.012)*	*(0.016)*
Siblings	-0.002	0.001	-0.004**	0.004
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*
Age	-0.080***	-0.034***	-0.046***	-0.036***
	(0.004)	*(0.008)*	*(0.003)*	*(0.005)*
Certainty of Punishment	0.003	-0.025***	0.005**	-0.022***
	(0.003)	*(0.003)*	*(0.002)*	*(0.003)*
Children	-0.018**	0.002	-0.025***	0.010
	(0.007)	*(0.008)*	*(0.008)*	*(0.008)*
Family Crime	0.002	0.176***	0.010	0.145***
	(0.015)	*(0.016)*	*(0.011)*	*(0.015)*
Drug Use			-0.045***	0.230***
			(0.009)	*(0.011)*
Unemployment Rate	0.021***	0.012**	0.018***	0.012**
	(0.006)	*(0.006)*	*(0.005)*	*(0.005)*
Future Outlook Inventory	0.019*	-0.038***	0.017*	-0.025**
	(0.011)	*(0.012)*	*(0.009)*	*(0.012)*
Years of Crime	-0.007***	0.028***	-0.005**	0.021***
	(0.003)	*(0.003)*	*(0.002)*	*(0.003)*
Years of Education	0.006	-0.003		
	(0.004)	*(0.004)*		
Schools per Young Person	0.321***		0.121*	
	(0.071)		*(0.063)*	
Lagged Enrollment	0.189***			
	(0.012)			
Enrollment		0.079		
		(0.052)		

Table A2: Average Marginal Effects from Probits for Crime and Education (Overall Crime) - Robustness Checks 2

Variable	**Excluding Drug Use**		**Enrollment Based on Attendance**	
	(1)		(2)	
	Educ.	Crime	Educ.	Crime
Lagged Crime		0.194***		0.156***
		(0.013)		*(0.013)*
Cognitive Factor	0.037	0.058**	-0.001	0.018
	(0.023)	*(0.025)*	*(0.018)*	*(0.024)*
Social/Emotional Factor	0.007	-0.122***	0.009	-0.084***
	(0.014)	*(0.015)*	*(0.011)*	*(0.015)*
Years of Education (alternative)			0.016***	0.004
			(0.004)	*(0.004)*
Lagged Enrollment (alternative)			0.084***	
			(0.010)	
Enrolment (alternative)				0.098*
				(0.056)
Rho	-0.115		-0.228*	
	(0.104)		*(0.124)*	
Observations	5,190	5,190	5,097	5,097

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. In column (1) we do not include drug use as an independent regressor. In column (2), enrollment is redefined as attending school for at least nine months in a year.

Table A3: Average Marginal Effects from Probits for Crime and Education (Overall Crime) - Robustness Checks 3

Variable	Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only	
	(1)		(2)		(3)		(4)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.051**	0.042**	0.048**	0.040**	0.048**	0.042**	0.053**	0.038*
	(0.021)	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.020)*
Hispanic	-0.025*	-0.021	-0.025*	-0.019	-0.025*	-0.019	-0.024*	-0.025
	(0.015)	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.015)*
Black	0.023	-0.030	0.025	-0.029	0.024	-0.030	0.018	-0.026
	(0.017)	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*
Other	0.035	-0.024	0.035	-0.023	0.032	-0.024	0.016	-0.011
	(0.027)	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*
Female	0.059***	-0.099***	0.058***	-0.101***	0.058***	-0.100***	0.061***	-0.118***
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.014)*	*(0.016)*
Non-intact Family	-0.052***	0.029*	-0.050***	0.030*	-0.051***	0.030*	-0.043***	0.027*
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.014)*	*(0.016)*
Siblings	-0.002	0.003	-0.002	0.004	-0.002	0.004	-0.002	0.004
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*
Age	-0.087***	-0.028***	-0.080***	-0.029***	-0.081***	-0.029***	-0.088***	-0.036***
	(0.005)	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.009)*
Certainty of Punishment	0.003	-0.022***	0.003	-0.022***	0.003	-0.022***	0.003	-0.022***
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.017**	0.007	-0.018**	0.008	-0.018**	0.008	-0.021***	0.007
	(0.007)	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.008)*
Family Crime	0.002	0.149***	0.002	0.148***	0.003	0.149***	-0.001	0.150***
	(0.015)	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.015)*
Drug Use	-0.000	0.224***	-0.001	0.224***	-0.001	0.224***	-0.002	0.223***
	(0.012)	*(0.010)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.010)*	*(0.011)*	*(0.011)*
Unemployment Rate	0.021***	0.011**	0.021***	0.011*	0.021***	0.011**	0.022***	0.011**
	(0.006)	*(0.005)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*
Future Outlook Inventory	0.020*	-0.023**	0.019*	-0.023**	0.019*	-0.024**	0.023**	-0.021*
	(0.011)	*(0.012)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*
Years of Crime			-0.015	0.008				
			(0.010)	*(0.010)*				
Years of Education			0.006	-0.002	0.007	-0.002	0.014***	-0.004
			(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.005)*
Schools per Young Person	0.318***		0.323***		0.325***		0.254***	
	(0.072)		*(0.071)*		*(0.071)*		*(0.070)*	
Lagged Enrollment			0.189***		0.190***		0.166***	
			(0.012)		*(0.012)*		*(0.013)*	
Enrollment				0.087*		0.081*		0.112**
				(0.049)		*(0.049)*		*(0.055)*
Lagged Crime				0.156***		0.158***		0.131***
				(0.013)		*(0.013)*		*(0.014)*
Cognitive Factor	0.036	0.016	0.035	0.012	0.036	0.014	0.038*	0.013
	(0.023)	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*
Social/Emotional Factor	0.006	-0.080***	0.007	-0.081***	0.006	-0.080***	0.009	-0.082***
	(0.014)	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*
Years of Crime * Age1	-0.006	0.022***						
	(0.004)	*(0.004)*						

Table A3: Average Marginal Effects from Probits for Crime and Education (Overall Crime) - Robustness Checks 3

Variable	Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only	
	(1)		(2)		(3)		(4)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Crime * Age2	-0.007**	0.020***						
	(0.003)	*(0.003)*						
Years of Education * Age1	0.002	-0.001						
	(0.005)	*(0.005)*						
Years of Education * Age2	0.009**	-0.002						
	(0.004)	*(0.005)*						
Lagged Enrollment * Age1	0.229***							
	(0.020)							
Lagged Enrollment * Age2	0.170***							
	(0.016)							
Enrollment * Age1		0.065						
		(0.050)						
Enrollment * Age2		0.046						
		(0.057)						
Lagged Crime * Age1		0.156***						
		(0.019)						
Lagged Crime * Age2		0.160***						
		(0.017)						
Years of Crime Squared			0.001	0.001				
			(0.001)	*(0.001)*				
Years of Crime: 0 to 4					-0.017**	0.023***		
					(0.008)	*(0.008)*		
Years of Crime: 5 to 9					-0.012***	0.021***		
					(0.004)	*(0.004)*		
Years of Crime: 10 or more					-0.008**	0.021***		
					(0.003)	*(0.003)*		
Years of Crime * Age of Entry 14							-0.051***	0.027***
							(0.006)	*(0.007)*
Years of Crime * Age of Entry 15							-0.026***	0.043***
							(0.005)	*(0.006)*
Years of Crime * Age of Entry 16							-0.022***	0.042***
							(0.006)	*(0.006)*
Years of Crime * Age of Entry 17							0.002	0.053***
							(0.006)	*(0.007)*
Years of Crime * Age of Entry 18							0.001	0.062***
							(0.012)	*(0.012)*
Rho	-0.074		-0.142		-0.127		-0.195	
	(0.111)		*(0.107)*		*(0.106)*		*(0.122)*	
Observations	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. In column (1) the coefficients are allowed to vary by age. Age1 is a dummy for ages 14 to 19, and Age2 is a dummy for ages 20 and above. In column (2) we use a quadratic function in criminal experience. In column (3) we use a piecewise-linear function of criminal experience: 0 to 4, 5 to 9, 10 or more. In column (4) we use the criminal experience observed in the sample only, interacted with age of entry dummies.

Table A4: Average Marginal Effects from Probits for Crime and Education (Overall Crime) - Robustness Checks 4

Variable	Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)	
	Educ.	Crime	Educ.	Crime
Phoenix	0.043**	0.051***	0.041*	-0.062
	(0.021)	*(0.019)*	*(0.021)*	*(0.062)*
Hispanic	-0.023	-0.023	-0.023	-0.023
	(0.015)	*(0.015)*	*(0.015)*	*(0.015)*
Black	0.027	-0.028	0.027	-0.028
	(0.017)	*(0.018)*	*(0.017)*	*(0.018)*
Other	0.035	-0.025	0.035	-0.028
	(0.027)	*(0.030)*	*(0.027)*	*(0.030)*
Female	0.067***	-0.098***	0.069***	-0.097***
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*
Non-intact Family	-0.053***	0.026*	-0.054***	0.026*
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*
Siblings	-0.002	0.003	-0.002	0.003
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*
Age	-0.074***	-0.040***	-0.072***	-0.036***
	(0.006)	*(0.004)*	*(0.006)*	*(0.004)*
Certainty of Punishment	0.005*	-0.022***	0.006**	-0.021***
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.018**	0.006	-0.018**	0.007
	(0.007)	*(0.007)*	*(0.007)*	*(0.007)*
Family Crime	-0.014	0.149***	-0.018	0.148***
	(0.018)	*(0.015)*	*(0.017)*	*(0.015)*
Drug Use	-0.029	0.225***	-0.035*	0.225***
	(0.020)	*(0.010)*	*(0.019)*	*(0.010)*
Unemployment Rate	0.019***	0.015***	0.019***	0.022***
	(0.006)	*(0.005)*	*(0.006)*	*(0.006)*
Future Outlook Inventory	0.021*	-0.022*	0.022**	-0.022*
	(0.011)	*(0.011)*	*(0.011)*	*(0.011)*
Years of Crime	-0.010***	0.020***	-0.011***	0.020***
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*
Years of Education	0.007	-0.001	0.007*	-0.000
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*
Schools per Young Person	0.301***		0.296***	
	(0.071)		*(0.071)*	
Lagged Enrollment	0.187***		0.186***	
	(0.012)		*(0.012)*	
Lagged Crime		0.160***		0.159***
		(0.012)		*(0.012)*
Cognitive Factor	0.036	0.016	0.036	0.017
	(0.023)	*(0.024)*	*(0.023)*	*(0.024)*
Social/Emotional Factor	0.015	-0.079***	0.017	-0.078***
	(0.015)	*(0.015)*	*(0.015)*	*(0.015)*
Crime	0.097*		0.120**	
	(0.058)		*(0.056)*	
Lagged State Arrest Rate				-1.815*
				(0.950)
Rho	-0.244		-0.186	
	(0.138)		*(0.141)*	
Observations	5,190	5,190	5,190	5,190

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. In column (1) we change the direction of the contemporaneous effect; we estimate the contemporaneous effect of crime on education. In column (2) we add the lagged state arrest rate as an exclusion in the crime equation.

Appendix B: Factor Model for Skills

Identification of the measurement/skills model of equations (4) and (5) follows from the analysis in Carneiro, Hansen, and Heckman (2003) and Cooley, Navarro, and Takahashi (2015). The argument roughly follows from first (conditionally) demeaning the measurements, which recovers the $\beta's$. The loadings (i.e., the δ's) are then identified by taking covariances between different cognitive measures and between different social/emotional measures. The marginal distributions of θ_i^{cog} and $\{\xi_{j,i,t_i}^{cog}\}_{j=1}^{J}$, as well as those of θ_i^{emo} and $\left\{\left\{\xi_{k,i,t}^{emo}\right\}_{t=t_i}^{T_i}\right\}_{k=1}^{K}$ are non-parametrically identified from a theorem of Kotlarski (1967) using deconvolution arguments. The correlation between θ_i^{cog} and θ_i^{emo} follows directly from the covariance between cognitive and social/emotional measures.

The distributions of the unobservables in the measurement systems are non-parametrically identified from the argument above. However, for estimation purposes, we impose distributional assumptions. In particular, we assume that $\xi_{j,i,t_i}^{cog} \sim N\left(0, \sigma_{\xi,cog,j}^2\right)$, $\xi_{k,i,t}^{emo} \sim N\left(0, \sigma_{\xi,emo,k}^2\right)$, and

$$\begin{pmatrix} \theta_i^{cog} \\ \theta_i^{emo} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\theta,cog}^2 & \rho\sigma_{\theta,cog}\sigma_{\theta,emo} \\ & \sigma_{\theta,emo}^2 \end{pmatrix}\right).$$

Given these distributional assumptions, the factor model is estimated by maximum likelihood. Let $\mathcal{M}_{j,i,t_i}^{cog} = M_{j,i,t_i}^{cog} - x_{i,t_i}\beta_j^{cog} - \theta_i^{cog}\delta_{j,t_i}^{cog}$, $\tilde{\psi}_{j,\ell} = \psi_{j,\ell} - x_{i,t_i}\beta_j^{cog} - \theta_i^{cog}\delta_{j,t_i}^{cog}$, and $\mathcal{M}_{k,i,t}^{emo} = M_{k,i,t}^{emo} - x_{i,t}\beta_{k,t}^{emo} - \theta_i^{emo}\delta_{k,t}^{emo}$. We define the conditional (on $\theta_i^{cog}, \theta_i^{emo}$) likelihood for the vector of individual observed test scores, M_i, to be

$$\begin{aligned} f\left(M_i|x_i, \theta_i^{cog}, \theta_i^{emo}; \beta, \psi, \delta, \sigma, \rho\right) = & \prod_{j=1}^{J_1} \phi\left(\mathcal{M}_{j,i,t_i}^{cog}|x_i, \theta_i^{cog}; \sigma_{\xi,cog,j}^2\right) \times \\ & \prod_{t=t_i}^{T_i}\prod_{k=1}^{K} \phi\left(\mathcal{M}_{k,i,t}^{emo}|x_i, \theta_i^{emo}; \sigma_{\xi,emo,k}^2\right) \times \\ & \prod_{j=J_1+1}^{J}\prod_{\ell=1}^{L_j} \left[\begin{array}{c} \Phi\left(\tilde{\psi}_{j,\ell}|x_i, \theta_i^{cog}; \beta_j^{cog}, \delta_j^{cog}\right) \\ -\Phi\left(\tilde{\psi}_{j,\ell-1}|x_i, \theta_i^{cog}; \beta_j^{cog}, \delta_j^{cog}\right) \end{array}\right] \mathbb{1}\left(M_{j,i} = \ell\right), \end{aligned}$$

where J_1 denotes the number of continuous cognitive tests, $J - J_1$ is the number of discrete tests, $\phi\left(|; \sigma^2\right)$ is the pdf of a mean zero normal with variance σ^2, and $\Phi\left(\right)$ is the cdf of a standard normal. The contribution to the likelihood of observation i is thus given by

$$f\left(M_i|x_i; \beta, \psi, \delta, \sigma, \rho\right) =$$

$$\int\int f\left(M_i|x_i,\theta_i^{cog},\theta_i^{emo};\beta,\psi,\delta,\sigma,\rho\right)\varphi\left(\theta_i^{cog},\theta_i^{emo};\sigma_{\theta,cog}^2,\sigma_{\theta,emo}^2,\rho\right)d\theta_i^{cog}d\theta_i^{emo},$$

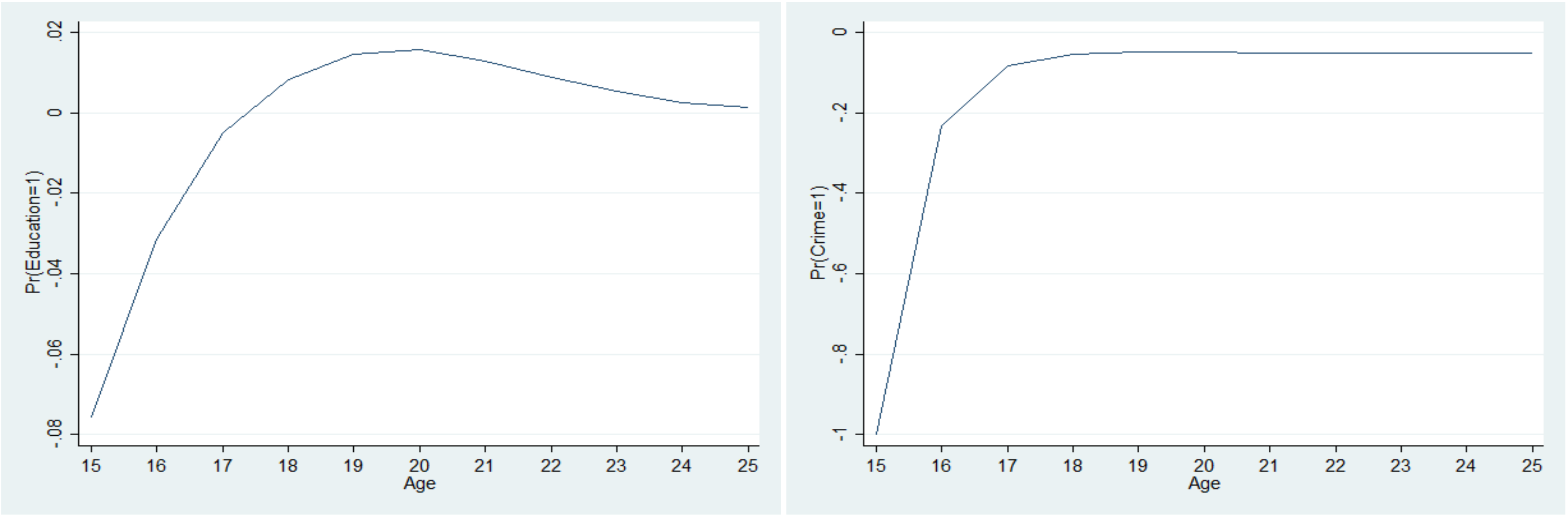
where $\varphi\left(;a,b,c\right)$ is the pdf of a mean zero bivariate normal with variances given by a,b and correlation coefficient c.

Having obtained estimates of the parameters of the factor model, we then predict the most likely values for $\theta_i^{cog},\theta_i^{emo}$ given the data we observe for each individual i. Prediction follows by applying Bayes' Rule to recover the distribution of $\theta_i^{cog},\theta_i^{emo}$ conditional on the data and then using it to obtain the expected value of $\theta_i^{cog},\theta_i^{emo}$ over that distribution. That is, we calculate

$$\begin{aligned}\begin{pmatrix}\bar{\theta}_i^{cog}\\ \bar{\theta}_i^{emo}\end{pmatrix} &= \int\int\begin{pmatrix}\theta_i^{cog}\\ \theta_i^{emo}\end{pmatrix} f\left(\theta_i^{cog},\theta_i^{emo}|M_i,x_i;\hat{\beta},\hat{\psi},\hat{\delta},\hat{\sigma},\hat{\rho}\right)d\theta_i^{cog}d\theta_i^{emo}\\ &= \int\int\begin{pmatrix}\theta_i^{cog}\\ \theta_i^{emo}\end{pmatrix}\frac{f\left(M_i|x_i,\theta_i^{cog},\theta_i^{emo};\hat{\beta},\hat{\psi},\hat{\delta},\hat{\sigma},\hat{\rho}\right)\varphi\left(\theta_i^{cog},\theta_i^{emo};\hat{\sigma}_{\theta,cog}^2,\hat{\sigma}_{\theta,emo}^2,\hat{\rho}\right)}{f\left(M_i|x_i;\hat{\beta},\hat{\psi},\hat{\delta},\hat{\sigma},\hat{\rho}\right)}d\theta_i^{cog}d\theta_i^{emo}.\end{aligned}$$

Online Appendix

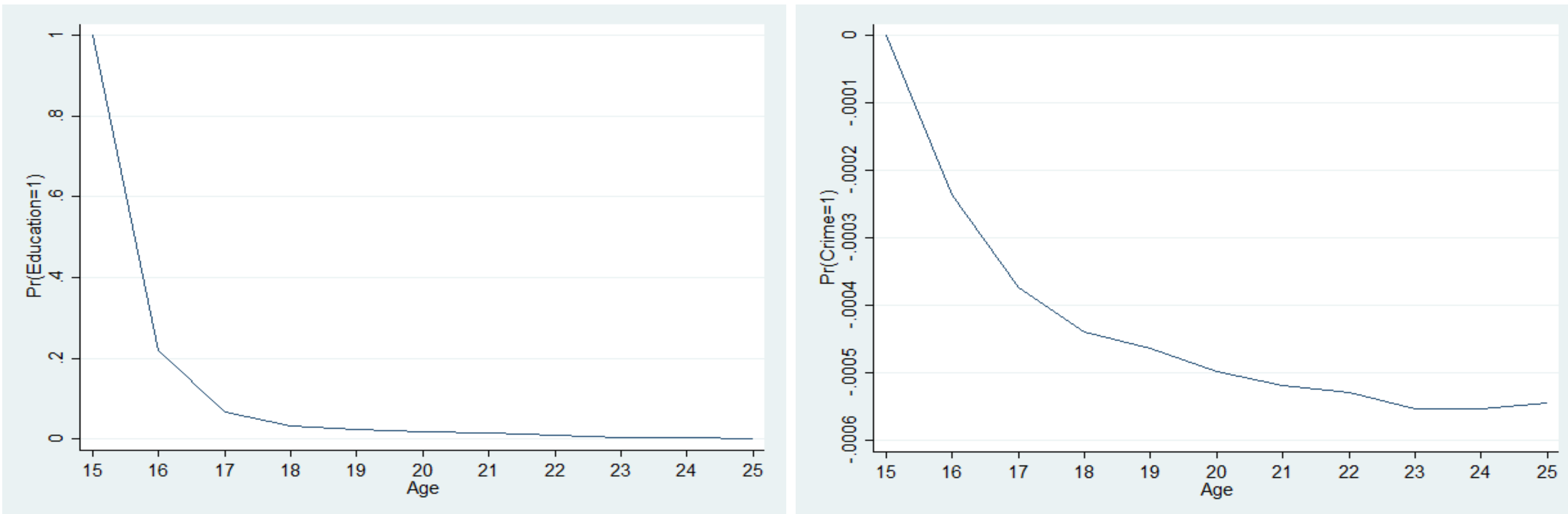
Figure O1: No Crime at Age 15 - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.
3. Note that for the second figure, the comparison between two identical individuals who differ only along one dimension (crime) implies that the average difference in the probability of crime between them is equal to -1 at age 15 by construction.

Figure O2: Enrolled at Age 15 - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



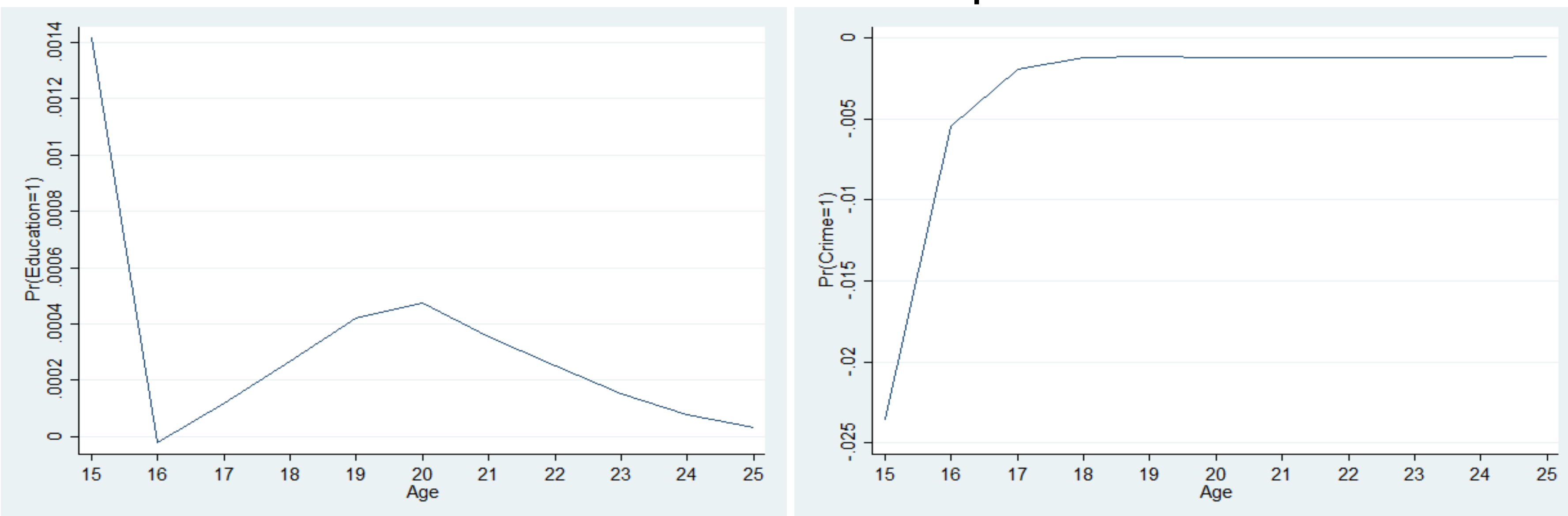
Notes:

1. The figures are based on the overall crime category.

2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

3. Note that for the first figure, the comparison between two identical individuals who differ only along one dimension (enrollment) at age 15 implies that the average difference in the probability of enrollment between them is equal to -1 at that age by construction.

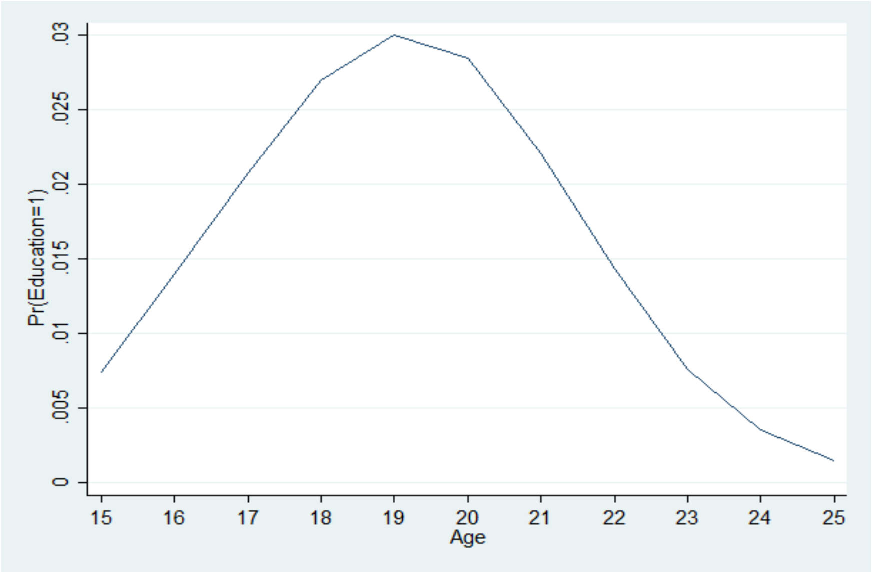
Figure O3: Increase in Certainty of Punishment at Age 15 - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect

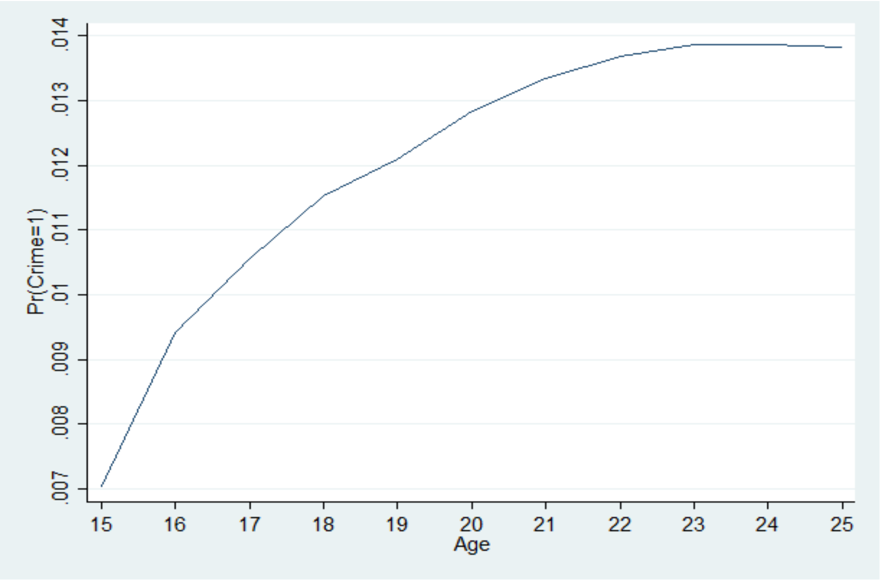


Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

Figure O4: Cognitive Factor 25th versus 75th Percentile - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



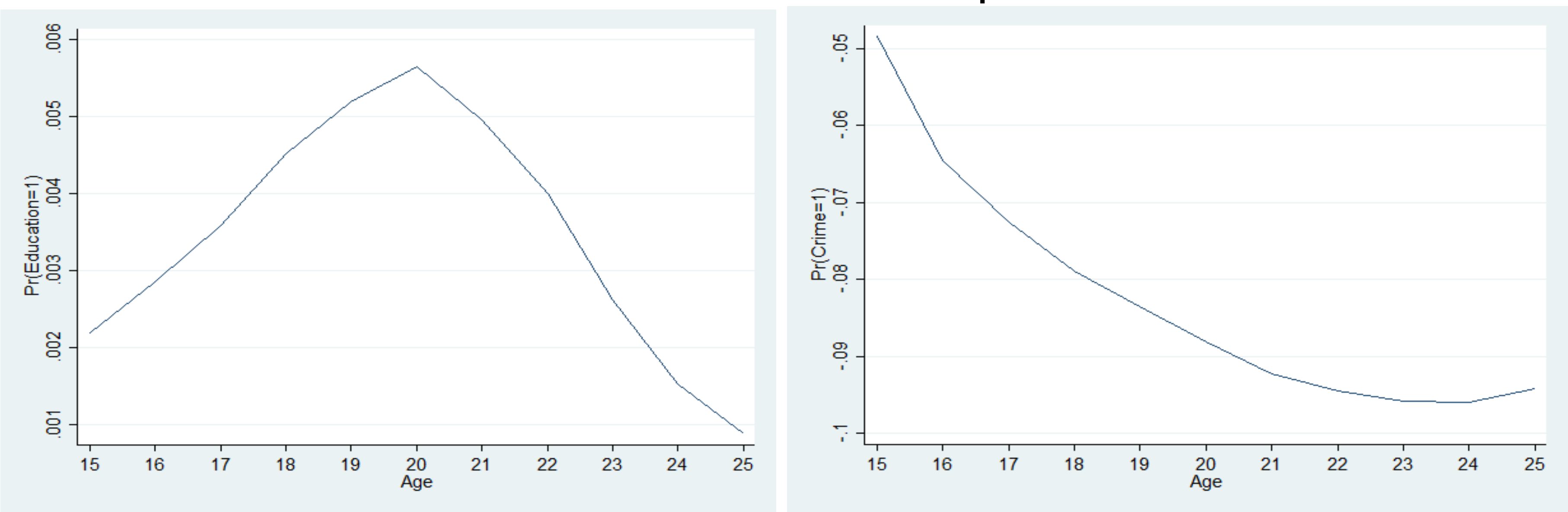


Notes:

1. The figures are based on the overall crime category.

2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

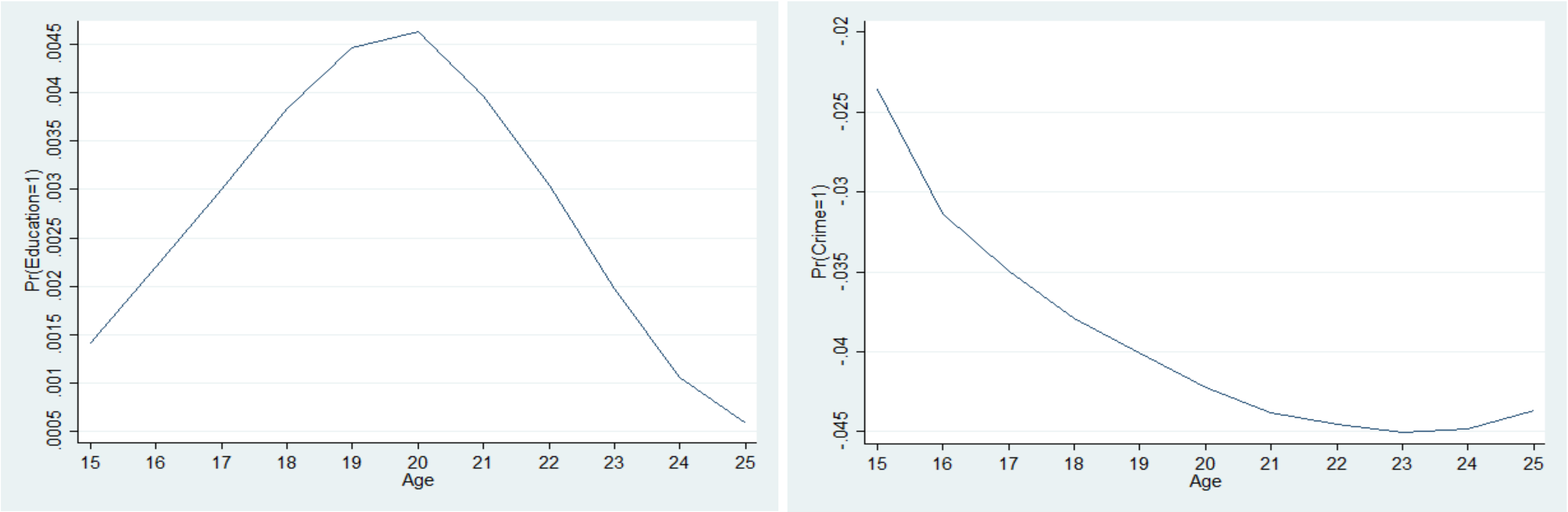
Figure O5: Social/Emotional Factor 25th versus 75th Percentile - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

Figure O6: Increase in Certainty of Punishment (Permanent) - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



Notes:

1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are then computed using the estimated parameters from the baseline model, and updated sequentially over time for a period of 10 years. We do this for 500,000 artificial agents and compute the average crime and enrollment rates.

Table O1: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime)

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.046**	-0.014	0.035*	-0.017	0.047**	-0.005	0.047**	-0.005	0.093***	-0.012	0.033	-0.014
	(0.021)	*(0.015)*	*(0.019)*	*(0.014)*	*(0.021)*	*(0.015)*	*(0.022)*	*(0.016)*	*(0.018)*	*(0.015)*	*(0.022)*	*(0.016)*
Hispanic	-0.025*	-0.022*			-0.023	-0.024**	-0.030**	-0.028**	-0.025*	-0.022*	-0.011	-0.027**
	(0.015)	*(0.012)*			*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*
Black	0.026	-0.007			0.027	-0.004	0.045**	-0.022	0.026	-0.006	0.043**	-0.012
	(0.017)	*(0.014)*			*(0.017)*	*(0.014)*	*(0.018)*	*(0.015)*	*(0.017)*	*(0.014)*	*(0.018)*	*(0.014)*
Other	0.034	-0.015			0.036	-0.011	0.039	-0.022	0.036	-0.014	0.042	-0.010
	(0.027)	*(0.024)*			*(0.028)*	*(0.024)*	*(0.029)*	*(0.025)*	*(0.028)*	*(0.025)*	*(0.028)*	*(0.025)*
Female	0.060***	-0.103***	0.056***	-0.099***	0.059***	-0.100***	0.067***	-0.142***	0.060***	-0.103***	0.053***	-0.099***
	(0.014)	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*
Non-intact Family	-0.053***	0.033***			-0.053***	0.029**	-0.055***	0.038***	-0.054***	0.032**	-0.051***	0.033***
	(0.015)	*(0.013)*			*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Siblings	-0.002	-0.000			-0.002	-0.000	-0.003	0.001	-0.002	-0.000	-0.002	0.000
	(0.002)	*(0.002)*			*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.081***	-0.000	-0.084***	-0.009	-0.081***	-0.010***	-0.104***	0.009*	-0.083***	-0.002	-0.080***	0.001
	(0.004)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.003)*	*(0.002)*	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.005)*
Certainty of Punishment	0.003	-0.010***			0.003	-0.010***	0.005**	-0.015***	0.003	-0.010***	0.003	-0.008***
	(0.003)	*(0.002)*			*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Children	-0.016**	0.009			-0.016**	0.007	-0.033***	0.019***	-0.016**	0.008	-0.016**	0.007
	(0.008)	*(0.006)*			*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*
Family Crime	0.001	0.084***			0.001	0.085***	-0.005	0.107***	0.003	0.084***	-0.000	0.082***
	(0.015)	*(0.010)*			*(0.015)*	*(0.010)*	*(0.015)*	*(0.011)*	*(0.015)*	*(0.010)*	*(0.015)*	*(0.010)*
Drug Use	0.006	0.213***			0.005	0.214***	-0.005	0.255***	0.007	0.213***	0.006	0.203***
	(0.012)	*(0.010)*			*(0.012)*	*(0.010)*	*(0.012)*	*(0.009)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.010)*
Unemployment Rate	0.020***	0.002	0.021***	0.002	0.021***	0.006	0.022***	0.001	0.037***	0.003	0.021***	0.001
	(0.006)	*(0.004)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.005)*	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Future Outlook Inventory	0.018	-0.008			0.019*	-0.006	0.022*	-0.008	0.016	-0.008	0.024*	0.007
	(0.011)	*(0.009)*			*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.012)*	*(0.010)*
Years of Crime	-0.011***	0.022***	-0.012***	0.040***	-0.011***	0.021***			-0.011***	0.022***	-0.011***	0.020***
	(0.004)	*(0.003)*	*(0.003)*	*(0.003)*	*(0.004)*	*(0.003)*			*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Years of Education	0.007	-0.008**	0.011***	-0.011***	0.007	-0.006*			0.006	-0.008**	0.006	-0.009***
	(0.004)	*(0.003)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.003)*			*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Cognitive Factor	0.049**	0.022			0.048**	0.026	0.061**	0.034*	0.048**	0.023		
	(0.023)	*(0.018)*			*(0.023)*	*(0.018)*	*(0.024)*	*(0.019)*	*(0.023)*	*(0.018)*		
Social/Emotional Factor	0.006	-0.028**			0.006	-0.028**	0.014	-0.049***	0.006	-0.028**		
	(0.014)	*(0.011)*			*(0.014)*	*(0.011)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.011)*		
Schools per Young Person	0.326***		0.324***		0.312***		0.323***				0.312***	
	(0.072)		*(0.072)*		*(0.072)*		*(0.073)*				*(0.071)*	
Lagged Enrollment	0.191***		0.194***		0.192***				0.192***		0.186***	
	(0.012)		*(0.012)*		*(0.012)*				*(0.013)*		*(0.013)*	
Enrollment		0.077**		0.063*		0.000		0.154***		0.062		0.088**
		(0.036)		*(0.038)*		*(0.011)*		*(0.039)*		*(0.038)*		*(0.034)*
Lagged Crime		0.097***		0.149***		0.098***				0.098***		0.092***
		(0.010)		*(0.011)*		*(0.010)*				*(0.010)*		*(0.010)*
WASI Reasoning Score											-0.001	0.001
											(0.006)	*(0.005)*
WASI Vocabulary Score											0.002	-0.002
											(0.007)	*(0.006)*
Stroop: Color/Word											0.005	-0.007
											(0.007)	*(0.005)*
Stroop: Word											0.010	-0.004
											(0.007)	*(0.006)*
Stroop: Color											-0.004	0.016**
											(0.008)	*(0.006)*

Table O1: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime)

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Trail-Making: Part B											-0.018**	0.004
											(0.007)	*(0.005)*
Trail-Making: Part A											-0.004	0.004
											(0.007)	*(0.005)*
WAI - Impulse Response											-0.008	-0.008
											(0.008)	*(0.006)*
WAI - Suppression of Aggression											0.011	-0.025***
											(0.007)	*(0.006)*
WAI - Consideration of Others											0.000	-0.012**
											(0.006)	*(0.005)*
PSMI - Self Reliance											-0.013	0.020**
											(0.010)	*(0.008)*
PSMI - Identity											0.036***	-0.029***
											(0.010)	*(0.008)*
PSMI - Work Orientation											-0.026***	0.010
											(0.009)	*(0.007)*
Rho	-0.289**		-0.198*				-0.548***		-0.230*		-0.320**	
	(0.131)		(0.109)				(0.140)		(0.137)		*(0.129)*	
Observations	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. The errors in the enrollment and crime equations are allowed to be correlated in every specification, expect for specification (3). Rho denotes the correlation in errors.
3. Every specification includes an exclusion restriction that enters the education equation only (Schools per Young Person) except for the specification in column (5).
4. In column (2) we exclude the cognitive and social/emotional factors and control for just a few variables (location, gender, age, and local unemployment rate). In column (3) the errors in the enrollment and crime equations are uncorrelated. The specification in column (4) does not account for any dynamics in the crime and education equation (years of experience and state dependence). The specification in column (5) does not include the exclusion restriction. In the last column we replace our factor estimates of cognitive and social/emotional skills with the measures used to infer them.

Table O2: Average Marginal Effects from Probits for Crime and Education (Violent Crime)

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.044**	0.037*	0.032*	0.036**	0.045**	0.045**	0.049**	0.032	0.092***	0.040**	0.035	0.043**
	(0.021)	*(0.020)*	*(0.019)*	*(0.018)*	*(0.021)*	*(0.019)*	*(0.021)*	*(0.021)*	*(0.018)*	*(0.020)*	*(0.022)*	*(0.021)*
Hispanic	-0.023	-0.018			-0.023	-0.020	-0.027*	-0.021	-0.024*	-0.019	-0.010	-0.026
	(0.015)	*(0.015)*			*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*
Black	0.027	-0.032*			0.027	-0.030	0.043**	-0.040**	0.027	-0.031*	0.043**	-0.032*
	(0.017)	*(0.018)*			*(0.017)*	*(0.018)*	*(0.018)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.018)*	*(0.019)*
Other	0.038	-0.015			0.037	-0.013	0.040	-0.011	0.039	-0.014	0.046*	-0.007
	(0.027)	*(0.030)*			*(0.027)*	*(0.030)*	*(0.028)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*
Female	0.056***	-0.079***	0.052***	-0.071***	0.056***	-0.076***	0.070***	-0.148***	0.055***	-0.078***	0.051***	-0.075***
	(0.015)	*(0.017)*	*(0.014)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*
Non-intact Family	-0.048***	0.014			-0.048***	0.010	-0.051***	0.021	-0.049***	0.013	-0.046***	0.015
	(0.015)	*(0.016)*			*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*
Siblings	-0.002	0.004			-0.002	0.004	-0.003	0.006**	-0.002	0.004	-0.002	0.004
	(0.002)	*(0.003)*			*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*
Age	-0.080***	-0.030***	-0.084***	-0.032***	-0.080***	-0.039***	-0.104***	-0.019**	-0.082***	-0.032***	-0.079***	-0.029***
	(0.004)	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.004)*	*(0.002)*	*(0.009)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.007)*
Certainty of Punishment	0.003	-0.019***			0.003	-0.019***	0.005*	-0.024***	0.003	-0.019***	0.002	-0.016***
	(0.003)	*(0.003)*			*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.017**	0.008			-0.017**	0.007	-0.032***	0.010	-0.017**	0.008	-0.017**	0.003
	(0.007)	*(0.008)*			*(0.007)*	*(0.008)*	*(0.008)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.007)*
Family Crime	0.005	0.130***			0.005	0.132***	-0.001	0.154***	0.006	0.130***	0.004	0.126***
	(0.015)	*(0.015)*			*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.014)*
Drug Use	-0.004	0.159***			-0.004	0.159***	-0.013	0.190***	-0.004	0.159***	-0.004	0.138***
	(0.011)	*(0.011)*			*(0.011)*	*(0.011)*	*(0.012)*	*(0.012)*	*(0.011)*	*(0.011)*	*(0.012)*	*(0.011)*
Unemployment Rate	0.021***	0.005	0.021***	0.004	0.021***	0.009*	0.022***	0.004	0.037***	0.006	0.021***	0.005
	(0.006)	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.006)*
Future Outlook Inventory	0.018*	-0.024**			0.017	-0.024**	0.025**	-0.028**	0.016	-0.024**	0.022*	0.012
	(0.011)	*(0.012)*			*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.012)*	*(0.013)*
Years of Crime	-0.007***	0.021***	-0.008***	0.033***	-0.007***	0.021***			-0.007***	0.021***	-0.007***	0.017***
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*			*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*
Years of Education	0.007	0.001	0.012***	-0.010**	0.006	0.003			0.006	0.002	0.006	-0.000
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*			*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Cognitive Factor	0.043*	0.030			0.045**	0.034	0.048**	0.054**	0.042*	0.032		
	(0.023)	*(0.024)*			*(0.023)*	*(0.024)*	*(0.024)*	*(0.024)*	*(0.023)*	*(0.024)*		
Social/Emotional Factor	0.007	-0.074***			0.008	-0.074***	0.018	-0.111***	0.008	-0.074***		
	(0.014)	*(0.015)*			*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.015)*		
Schools per Young Person	0.332***		0.334***		0.322***		0.326***				0.319***	
	(0.071)		*(0.071)*		*(0.071)*		*(0.073)*				*(0.071)*	
Lagged Enrollment	0.189***		0.191***		0.190***				0.190***		0.185***	
	(0.012)		*(0.012)*		*(0.012)*				*(0.012)*		*(0.012)*	
Enrollment		0.104**		0.101**		0.033**		0.199***		0.083		0.106**
		(0.048)		*(0.051)*		*(0.014)*		*(0.061)*		*(0.051)*		*(0.047)*
Lagged Crime		0.142***		0.188***		0.144***				0.143***		0.125***
		(0.012)		*(0.013)*		*(0.012)*				*(0.012)*		*(0.012)*
WASI Reasoning Score											0.000	-0.005
											(0.006)	*(0.007)*
WASI Vocabulary Score											-0.002	-0.009
											(0.007)	*(0.007)*
Stroop: Color/Word											0.004	-0.002
											(0.007)	*(0.007)*
Stroop: Word											0.010	-0.003
											(0.007)	*(0.008)*
Stroop: Color											-0.002	0.008
											(0.008)	*(0.008)*

Table O2: Average Marginal Effects from Probits for Crime and Education (Violent Crime)

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Trail-Making: Part B											-0.016**	-0.007
											(0.007)	*(0.007)*
Trail-Making: Part A											-0.003	-0.001
											(0.007)	*(0.007)*
WAI - Impulse Response											-0.009	-0.025***
											(0.007)	*(0.008)*
WAI - Suppression of Aggression											0.011	-0.059***
											(0.007)	*(0.007)*
WAI - Consideration of Others											0.001	-0.027***
											(0.006)	*(0.006)*
PSMI - Self Reliance											-0.014	0.013
											(0.010)	*(0.011)*
PSMI - Identity											0.035***	-0.008
											(0.010)	*(0.011)*
PSMI - Work Orientation											-0.022**	0.004
											(0.009)	*(0.009)*
Rho	-0.156		-0.157				-0.348**		-0.109		-0.159	
	(0.104)		(0.102)				(0.137)		(0.108)		*(0.104)*	
Observations	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. The errors in the enrollment and crime equations are allowed to be correlated in every specification, expect for specification (3). Rho denotes the correlation in errors.
3. Every specification includes an exclusion restriction that enters the education equation only (Schools per Young Person) except for the specification in column (5).
4. In column (2) we exclude the cognitive and social/emotional factors and control for just a few variables (location, gender, age, and local unemployment rate). In column (3) the errors in the enrollment and crime equations are uncorrelated. The specification in column (4) does not account for any dynamics in the crime and education equation (years of experience and state dependence). The specification in column (5) does not include the exclusion restriction. In the last column we replace our factor estimates of cognitive and social/emotional skills with the measures used to infer them.

Table O3: Average Marginal Effects from Probits for Crime and Education (Property Crime)

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.054**	0.062***	0.046**	0.029*	0.054**	0.059***	0.053**	0.124***	0.102***	0.063***	0.046**	0.046***
	(0.021)	*(0.017)*	*(0.019)*	*(0.016)*	*(0.021)*	*(0.017)*	*(0.022)*	*(0.019)*	*(0.018)*	*(0.017)*	*(0.022)*	*(0.017)*
Hispanic	-0.024	-0.012			-0.024	-0.012	-0.028*	-0.037***	-0.025*	-0.013	-0.013	-0.011
	(0.015)	*(0.013)*			*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Black	0.024	-0.006			0.024	-0.006	0.044**	-0.029*	0.024	-0.005	0.039**	-0.001
	(0.017)	*(0.015)*			*(0.017)*	*(0.015)*	*(0.018)*	*(0.016)*	*(0.017)*	*(0.015)*	*(0.018)*	*(0.016)*
Other	0.035	0.015			0.035	0.015	0.040	-0.005	0.036	0.016	0.042	0.021
	(0.027)	*(0.025)*			*(0.027)*	*(0.025)*	*(0.028)*	*(0.026)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*
Female	0.066***	-0.024*	0.062***	-0.029**	0.067***	-0.026*	0.069***	-0.056***	0.066***	-0.024*	0.060***	-0.021
	(0.014)	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*
Non-intact Family	-0.048***	0.007			-0.048***	0.008	-0.050***	0.008	-0.049***	0.007	-0.046***	0.005
	(0.015)	*(0.013)*			*(0.015)*	*(0.013)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Siblings	-0.003	0.002			-0.003	0.002	-0.004	0.002	-0.003	0.002	-0.003	0.002
	(0.002)	*(0.002)*			*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.082***	-0.027***	-0.086***	-0.030***	-0.082***	-0.024***	-0.104***	-0.028***	-0.084***	-0.029***	-0.081***	-0.025***
	(0.004)	*(0.006)*	*(0.004)*	*(0.007)*	*(0.004)*	*(0.004)*	*(0.002)*	*(0.010)*	*(0.004)*	*(0.007)*	*(0.004)*	*(0.006)*
Certainty of Punishment	0.004	-0.016***			0.004	-0.016***	0.005**	-0.022***	0.003	-0.016***	0.003	-0.013***
	(0.003)	*(0.002)*			*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Children	-0.017**	0.001			-0.018**	0.001	-0.032***	0.003	-0.017**	0.001	-0.017**	-0.002
	(0.007)	*(0.007)*			*(0.007)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Family Crime	0.001	0.095***			0.001	0.095***	-0.001	0.122***	0.003	0.095***	0.001	0.090***
	(0.015)	*(0.011)*			*(0.015)*	*(0.011)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.011)*	*(0.015)*	*(0.011)*
Drug Use	-0.007	0.144***			-0.006	0.144***	-0.012	0.181***	-0.006	0.144***	-0.005	0.126***
	(0.011)	*(0.009)*			*(0.011)*	*(0.009)*	*(0.012)*	*(0.010)*	*(0.011)*	*(0.009)*	*(0.012)*	*(0.009)*
Unemployment Rate	0.021***	0.010**	0.022***	0.008*	0.021***	0.009**	0.024***	0.012**	0.038***	0.010**	0.022***	0.008*
	(0.006)	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.005)*	*(0.006)*	*(0.005)*
Future Outlook Inventory	0.017	-0.032***			0.017	-0.033***	0.023**	-0.040***	0.015	-0.032***	0.022*	0.003
	(0.011)	*(0.010)*			*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.012)*	*(0.011)*
Years of Crime	-0.003	0.018***	-0.004*	0.030***	-0.003	0.018***			-0.003	0.018***	-0.003	0.015***
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*			*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Years of Education	0.007*	0.006*	0.012***	-0.001	0.007*	0.006			0.006	0.007*	0.007	0.003
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*			*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Cognitive Factor	0.041*	0.020			0.040*	0.019	0.052**	0.047**	0.040*	0.020		
	(0.023)	*(0.020)*			*(0.023)*	*(0.020)*	*(0.024)*	*(0.021)*	*(0.023)*	*(0.020)*		
Social/Emotional Factor	0.007	-0.068***			0.007	-0.068***	0.017	-0.127***	0.007	-0.068***		
	(0.014)	*(0.013)*			*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*		
Schools per Young Person	0.327***		0.328***		0.329***		0.310***				0.314***	
	(0.071)		*(0.072)*		*(0.071)*		*(0.074)*				*(0.071)*	
Lagged Enrollment	0.190***		0.192***		0.191***				0.191***		0.186***	
	(0.012)		*(0.012)*		*(0.012)*				*(0.012)*		*(0.012)*	
Enrollment		-0.022		-0.019		-0.003		-0.017		-0.035		-0.011
		(0.040)		*(0.041)*		*(0.012)*		*(0.077)*		*(0.042)*		*(0.040)*
Lagged Crime		0.144***		0.196***		0.144***				0.144***		0.132***
		(0.010)		*(0.011)*		*(0.010)*				*(0.010)*		*(0.010)*
WASI Reasoning Score											0.001	0.001
											(0.006)	*(0.006)*
WASI Vocabulary Score											-0.003	0.013**
											(0.007)	*(0.006)*
Stroop: Color/Word											0.003	-0.011*
											(0.007)	*(0.006)*
Stroop: Word											0.008	-0.002
											(0.007)	*(0.006)*
Stroop: Color											-0.001	0.005
											(0.008)	*(0.007)*

Table O3: Average Marginal Effects from Probits for Crime and Education (Property Crime)

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Trail-Making: Part B											-0.016**	-0.007
											(0.007)	*(0.006)*
Trail-Making: Part A											-0.004	0.000
											(0.007)	*(0.006)*
WAI - Impulse Response											-0.007	-0.030***
											(0.007)	*(0.006)*
WAI - Suppression of Aggression											0.013*	-0.029***
											(0.007)	*(0.006)*
WAI - Consideration of Others											0.001	-0.022***
											(0.006)	*(0.005)*
PSMI - Self Reliance											-0.013	0.012
											(0.010)	*(0.009)*
PSMI - Identity											0.035***	-0.012
											(0.010)	*(0.009)*
PSMI - Work Orientation											-0.024***	-0.015**
											(0.009)	*(0.008)*
Rho	0.061		0.014				0.039		0.010		0.041	
	(0.118)		(0.107)				(0.202)		(0.123)		*(0.119)*	
Observations	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. The errors in the enrollment and crime equations are allowed to be correlated in every specification, expect for specification (3). Rho denotes the correlation in errors.
3. Every specification includes an exclusion restriction that enters the education equation only (Schools per Young Person) except for the specification in column (5).
4. In column (2) we exclude the cognitive and social/emotional factors and control for just a few variables (location, gender, age, and local unemployment rate). In column (3) the errors in the enrollment and crime equations are uncorrelated. The specification in column (4) does not account for any dynamics in the crime and education equation (years of experience and state dependence). The specification in column (5) does not include the exclusion restriction. In the last column we replace our factor estimates of cognitive and social/emotional skills with the measures used to infer them.

Table O4: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.046**	-0.013	0.067***	0.001	0.052***	-0.005	0.051***	-0.007	0.014	-0.005	0.048**	-0.012	0.045**	-0.012	0.045**	-0.012	0.045**	-0.011	0.045**	-0.063	0.045**	-0.003
	(0.021)	*(0.016)*	*(0.021)*	*(0.014)*	*(0.019)*	*(0.014)*	*(0.019)*	*(0.014)*	*(0.018)*	*(0.015)*	*(0.021)*	*(0.015)*	*(0.021)*	*(0.015)*	*(0.021)*	*(0.015)*	*(0.021)*	*(0.015)*	*(0.021)*	*(0.047)*	*(0.021)*	*(0.015)*
Hispanic	-0.024*	-0.030**	-0.032**	-0.014	-0.005	-0.019*	-0.005	-0.019*	-0.020*	-0.023**	-0.024*	-0.021*	-0.024*	-0.022*	-0.024*	-0.022*	-0.024	-0.024**	-0.019	-0.023**	-0.019	-0.023*
	(0.015)	*(0.012)*	*(0.015)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.012)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*
Black	0.026	-0.020	-0.003	0.002	0.045***	-0.006	0.043***	-0.008	0.001	-0.007	0.025	-0.007	0.026	-0.006	0.025	-0.006	0.023	-0.011	0.026	-0.005	0.026	-0.004
	(0.017)	*(0.015)*	*(0.017)*	*(0.013)*	*(0.016)*	*(0.013)*	*(0.016)*	*(0.013)*	*(0.014)*	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*
Other	0.033	-0.036	0.012	-0.007	0.033	-0.009	0.032	-0.009	-0.022	-0.012	0.036	-0.014	0.034	-0.015	0.033	-0.014	0.027	-0.011	0.035	-0.015	0.035	-0.014
	(0.027)	*(0.026)*	*(0.027)*	*(0.023)*	*(0.027)*	*(0.023)*	*(0.027)*	*(0.023)*	*(0.022)*	*(0.025)*	*(0.027)*	*(0.024)*	*(0.028)*	*(0.025)*	*(0.028)*	*(0.024)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*
Female	0.060***	-0.103***	0.114***	-0.092***	0.023	-0.078***	0.024	-0.076***	0.013	-0.100***	0.062***	-0.101***	0.060***	-0.104***	0.060***	-0.103***	0.058***	-0.108***	0.070***	-0.100***	0.070***	-0.100***
	(0.014)	*(0.015)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*
Non-intact Family	-0.053***	0.035**	-0.062***	0.021*	-0.033**	0.019	-0.033**	0.019	-0.007	0.030**	-0.054***	0.033***	-0.053***	0.033***	-0.053***	0.033***	-0.051***	0.034***	-0.057***	0.028**	-0.057***	0.028**
	(0.015)	*(0.014)*	*(0.015)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.012)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Siblings	-0.002	-0.002	-0.002	-0.001	-0.002	-0.001	-0.002	-0.001	-0.004**	-0.001	-0.002	-0.000	-0.002	-0.000	-0.002	-0.000	-0.002	0.001	-0.002	-0.000	-0.002	-0.000
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.081***	-0.006	-0.069***	-0.010***	-0.076***	-0.004	-0.075***	-0.005	-0.048***	-0.013***	-0.088***	-0.001	-0.081***	-0.000	-0.081***	-0.000	-0.084***	-0.006	-0.078***	-0.008**	-0.078***	-0.010***
	(0.004)	*(0.006)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.005)*	*(0.003)*	*(0.004)*	*(0.005)*	*(0.006)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Certainty of Punishment	0.003	-0.013***	0.007***	-0.008***	-0.000	-0.007***	-0.000	-0.007***	0.006***	-0.010***	0.003	-0.010***	0.003	-0.010***	0.003	-0.010***	0.003	-0.011***	0.005*	-0.010***	0.005*	-0.010***
	(0.003)	*(0.002)*	*(0.003)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Children	-0.017**	0.004	-0.013	0.010*	-0.015**	0.013**	-0.013**	0.013**	-0.024***	0.012*	-0.016**	0.009	-0.016**	0.009	-0.016**	0.009	-0.019**	0.007	-0.018**	0.008	-0.018**	0.007
	(0.008)	*(0.006)*	*(0.008)*	*(0.005)*	*(0.007)*	*(0.005)*	*(0.007)*	*(0.005)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*
Family Crime	0.002	0.108***	-0.029**	0.078***	0.005	0.069***	0.004	0.069***	0.009	0.083***	0.002	0.083***	0.001	0.084***	0.000	0.085***	0.001	0.083***	-0.019	0.084***	-0.018	0.084***
	(0.015)	*(0.011)*	*(0.014)*	*(0.009)*	*(0.014)*	*(0.009)*	*(0.014)*	*(0.009)*	*(0.011)*	*(0.010)*	*(0.015)*	*(0.010)*	*(0.015)*	*(0.010)*	*(0.015)*	*(0.010)*	*(0.015)*	*(0.010)*	*(0.017)*	*(0.010)*	*(0.017)*	*(0.010)*
Drug Use			0.056***	0.228***	-0.038***	0.235***	-0.039***	0.235***	-0.044***	0.217***	0.007	0.212***	0.006	0.213***	0.006	0.213***	0.008	0.216***	-0.028	0.214***	-0.026	0.214***
			(0.011)	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.009)*	*(0.010)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.010)*	*(0.018)*	*(0.010)*	*(0.018)*	*(0.010)*
Unemployment Rate	0.020***	0.002	0.020***	0.007*	0.022***	0.004	0.022***	0.004	0.018***	0.006	0.020***	0.003	0.020***	0.002	0.020***	0.002	0.021***	0.004	0.019***	0.010**	0.019***	0.007*
	(0.006)	*(0.005)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.004)*
Future Outlook Inventory	0.018	-0.023**	0.019*	-0.009	0.015	-0.007	0.015	-0.008	0.017*	-0.006	0.020*	-0.008	0.018*	-0.008	0.018*	-0.008	0.021*	-0.005	0.020*	-0.007	0.020*	-0.007
	(0.011)	*(0.010)*	*(0.011)*	*(0.008)*	*(0.010)*	*(0.008)*	*(0.010)*	*(0.008)*	*(0.009)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*
Years of Crime	-0.010***	0.033***	-0.021***	0.023***	-0.002	0.021***	-0.003	0.023***	-0.003	0.021***			-0.017**	0.029***					-0.016***	0.021***	-0.016***	0.021***
	(0.004)	*(0.003)*	*(0.004)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.003)*	*(0.003)*			*(0.008)*	*(0.006)*					*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Years of Education	0.007	-0.006*	0.022***	-0.004	-0.007*	-0.003	-0.002	-0.003					0.007	-0.008**	0.007	-0.008**	0.010**	-0.007*	0.008*	-0.007**	0.008*	-0.007**
	(0.004)	*(0.004)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.004)*					*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Schools per Young Person	0.325***		0.146**		0.322***		0.324***		0.141**		0.332***		0.328***		0.327***		0.297***		0.304***		0.307***	
	(0.072)		*(0.072)*		*(0.067)*		*(0.067)*		*(0.065)*		*(0.072)*		*(0.072)*		*(0.072)*		*(0.072)*		*(0.072)*		*(0.071)*	
Lagged Enrollment	0.191***		0.179***		0.199***		0.223***						0.191***		0.191***		0.182***		0.188***		0.189***	
	(0.012)		*(0.013)*		*(0.011)*		*(0.013)*						*(0.012)*		*(0.012)*		*(0.013)*		*(0.013)*		*(0.013)*	
Enrollment		0.069*		-0.002										0.077**		0.075**		0.057				
		(0.038)		*(0.012)*										*(0.035)*		*(0.036)*		*(0.040)*				
Lagged Crime		0.130***		0.088***		0.085***		0.098***		0.098***				0.096***		0.096***		0.089***		0.096***		0.096***
		(0.011)		*(0.010)*		*(0.010)*		*(0.012)*		*(0.010)*				*(0.010)*		*(0.010)*		*(0.011)*		*(0.010)*		*(0.010)*
Cognitive Factor	0.050**	0.060***	0.042*	0.016	0.007	0.025	0.005	0.024	0.004	0.030*	0.047**	0.020	0.049**	0.023	0.049**	0.023	0.048**	0.022	0.043*	0.024	0.044*	0.024
	(0.023)	*(0.019)*	*(0.023)*	*(0.017)*	*(0.022)*	*(0.017)*	*(0.022)*	*(0.017)*	*(0.019)*	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*
Social/Emotional Factor	0.005	-0.066***	0.014	-0.026**	-0.011	-0.025**	-0.012	-0.025**	0.012	-0.031***	0.004	-0.029**	0.005	-0.028**	0.006	-0.028**	0.007	-0.025**	0.009	-0.024**	0.009	-0.025**
	(0.014)	*(0.012)*	*(0.013)*	*(0.011)*	*(0.013)*	*(0.010)*	*(0.013)*	*(0.010)*	*(0.011)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*
Jail					0.104***	0.063***	0.367***	0.068														
					(0.012)	*(0.010)*	*(0.076)*	*(0.065)*														
Enrollment (alternative)						0.042		0.038														
						(0.035)		*(0.035)*														

Table O4: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Crime * Jail							0.004	-0.006														
							(0.007)	*(0.006)*														
Years of Education * Jail							-0.019***	0.002														
							(0.007)	*(0.005)*														
Lagged Enrollment * Jail							-0.088***															
							(0.023)															
Lagged Crime * Jail								-0.034*														
								(0.020)														
Enrollment * Jail								-0.001														
								(0.019)														
Years of Education (alternative)									0.017***	-0.002												
									(0.004)	*(0.003)*												
Lagged Enrollment (alternative)									0.086***													
									(0.010)													
Enrolment (alternative)										-0.005												
										(0.048)												
Years of Crime * Age1											-0.016***	0.027***										
											(0.006)	*(0.005)*										
Years of Crime * Age2											-0.008*	0.019***										
											(0.004)	*(0.004)*										
Years of Education * Age1											0.004	-0.008**										
											(0.005)	*(0.004)*										
Years of Education * Age2											0.009**	-0.008**										
											(0.004)	*(0.003)*										
Lagged Enrollment * Age1											0.228***											
											(0.020)											
Lagged Enrollment * Age2											0.172***											
											(0.016)											
Enrollment * Age1												0.080**										
												(0.036)										
Enrollment * Age2												0.089**										
												(0.045)										
Lagged Crime * Age1												0.072***										
												(0.015)										
Lagged Crime * Age2												0.123***										
												(0.015)										
Years of Crime Squared													0.001	-0.001								
													(0.002)	*(0.001)*								
Years of Crime: 0 to 4															-0.013***	0.024***						
															(0.004)	*(0.004)*						
Years of Crime: 5 to 9															-0.006	0.019***						
															(0.005)	*(0.003)*						
Years of Crime: 10 or more															-0.088	0.119						
															(32.524)	*(23.080)*						
Years of Crime * Age of Entry 14																	-0.063***	0.019**				
																	(0.013)	*(0.009)*				

Table O4: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Crime * Age of Entry 15																	-0.035***	0.023***				
																	(0.011)	*(0.008)*				
Years of Crime * Age of Entry 16																	-0.023***	0.029***				
																	(0.008)	*(0.006)*				
Years of Crime * Age of Entry 17																	-0.000	0.046***				
																	(0.009)	*(0.006)*				
Years of Crime * Age of Entry 18																	-0.003	0.029***				
																	(0.019)	*(0.011)*				
Crime																			0.147**		0.137**	
																			(0.058)		*(0.062)*	
Lagged State Arrest Rate																				-0.962		
																				(0.727)		
Rho	-0.214*		0.030		-0.166		-0.152		-0.051		-0.311**		-0.287**		-0.281**		-0.225		-0.353**		-0.379**	
	(0.117)		*(0.041)*		*(0.121)*		*(0.121)*		*(0.165)*		*(0.142)*		*(0.130)*		*(0.131)*		*(0.145)*		*(0.162)*		*(0.155)*	
Observations	5,074	5,074	6,042	6,042	6,042	6,042	6,042	6,042	4,987	4,987	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5074	5074	5074	5074

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.

2. In column (1) we do not include drug use as an independent regressor. In column (2), enrollment is set to zero if an individual did not attend a community school. In column (3), we condition on whether the individual is interviewed in jail, and in column (4) we interact the jail dummy with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail. In column (5) enrollment is redefined as attending school for at least nine months. Coefficients are allowed to vary by age in specification (6). Age1 is a dummy for ages 14 to 19, and Age2 is a dummy for ages 20 and above. In column (7) we use a quadratic function in criminal experience. In column (8) we use a piecewise-linear function of criminal experience: 0 to 4, 5 to 9, more than 10. In column (9) we use the criminal experience observed in the sample only, interacted with age of entry dummies. In column (10) we change the direction of the contemporaneous effect; we estimate the contemporaneous effect of crime on education. In column (11) we add the lagged state arrest rate as an exclusion in the crime equation.

Table O5: Average Marginal Effects from Probits for Crime and Education (Violent Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.044**	0.038*	0.065***	0.041**	0.054***	0.036*	0.052***	0.035*	0.012	0.041**	0.047**	0.038*	0.045**	0.036*	0.041**	0.036*	0.047**	0.024	0.031	-0.048	0.031	0.048**
	(0.021)	*(0.021)*	*(0.021)*	*(0.018)*	*(0.019)*	*(0.019)*	*(0.019)*	*(0.019)*	*(0.018)*	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.062)*	*(0.021)*	*(0.019)*
Hispanic	-0.023	-0.023	-0.033**	-0.021	-0.005	-0.025*	-0.005	-0.026*	-0.018	-0.015	-0.023	-0.019	-0.023	-0.017	-0.021	-0.017	-0.022	-0.020	-0.019	-0.021	-0.019	-0.021
	(0.015)	*(0.016)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.012)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*
Black	0.027	-0.039**	0.003	-0.035**	0.043***	-0.041**	0.040**	-0.041**	0.007	-0.027	0.026	-0.031*	0.027	-0.031*	0.030*	-0.031*	0.021	-0.025	0.032*	-0.029	0.032*	-0.029
	(0.017)	*(0.019)*	*(0.017)*	*(0.017)*	*(0.016)*	*(0.017)*	*(0.016)*	*(0.017)*	*(0.013)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*
Other	0.038	-0.028	0.011	-0.005	0.034	-0.007	0.032	-0.007	-0.019	-0.010	0.038	-0.013	0.038	-0.014	0.040	-0.012	0.019	0.003	0.035	-0.019	0.036	-0.017
	(0.027)	*(0.031)*	*(0.027)*	*(0.028)*	*(0.026)*	*(0.028)*	*(0.026)*	*(0.028)*	*(0.022)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.031)*	*(0.027)*	*(0.031)*
Female	0.056***	-0.079***	0.116***	-0.078***	0.017	-0.059***	0.018	-0.057***	0.002	-0.070***	0.057***	-0.077***	0.056***	-0.079***	0.054***	-0.078***	0.063***	-0.101***	0.068***	-0.075***	0.067***	-0.076***
	(0.015)	*(0.017)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.017)*	*(0.012)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.014)*	*(0.016)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*
Non-intact Family	-0.048***	0.011	-0.055***	0.010	-0.028**	0.005	-0.029**	0.004	-0.003	0.010	-0.049***	0.012	-0.048***	0.014	-0.047***	0.014	-0.044***	0.017	-0.048***	0.009	-0.048***	0.009
	(0.015)	*(0.016)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.012)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.014)*	*(0.016)*	*(0.014)*	*(0.016)*	*(0.014)*	*(0.016)*
Siblings	-0.002	0.002	-0.003	0.003	-0.002	0.002	-0.001	0.002	-0.005**	0.004	-0.002	0.004	-0.002	0.004	-0.002	0.004	-0.002	0.004	-0.003	0.004	-0.003	0.004
	(0.002)	*(0.003)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*
Age	-0.080***	-0.033***	-0.071***	-0.036***	-0.074***	-0.038***	-0.074***	-0.039***	-0.045***	-0.036***	-0.087***	-0.031***	-0.080***	-0.030***	-0.081***	-0.030***	-0.089***	-0.034***	-0.066***	-0.041***	-0.068***	-0.044***
	(0.004)	*(0.008)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.007)*	*(0.004)*	*(0.007)*	*(0.003)*	*(0.005)*	*(0.005)*	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.009)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*
Certainty of Punishment	0.003	-0.022***	0.007***	-0.017***	-0.001	-0.015***	-0.001	-0.015***	0.006***	-0.019***	0.003	-0.019***	0.003	-0.019***	0.003	-0.020***	0.003	-0.020***	0.007***	-0.019***	0.007**	-0.019***
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.017**	0.004	-0.015**	0.008	-0.015**	0.011	-0.014**	0.010	-0.024***	0.007	-0.017**	0.007	-0.017**	0.008	-0.018**	0.008	-0.020***	0.008	-0.018**	0.006	-0.018**	0.006
	(0.007)	*(0.008)*	*(0.008)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.008)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*
Family Crime	0.004	0.150***	-0.032**	0.136***	0.011	0.126***	0.010	0.127***	0.012	0.127***	0.005	0.131***	0.005	0.130***	0.005	0.129***	0.001	0.130***	-0.029*	0.130***	-0.026	0.131***
	(0.015)	*(0.015)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.011)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*
Drug Use			0.041***	0.160***	-0.043***	0.164***	-0.044***	0.164***	-0.042***	0.165***	-0.004	0.159***	-0.004	0.159***	-0.005	0.159***	-0.007	0.161***	-0.042***	0.160***	-0.039***	0.160***
			(0.011)	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.009)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.015)*	*(0.011)*
Unemployment Rate	0.021***	0.006	0.019***	0.007	0.023***	0.007	0.023***	0.006	0.017***	0.006	0.021***	0.006	0.021***	0.005	0.020***	0.004	0.022***	0.004	0.018***	0.016***	0.018***	0.011**
	(0.006)	*(0.006)*	*(0.006)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.005)*
Future Outlook Inventory	0.018*	-0.034***	0.018*	-0.031***	0.015	-0.029***	0.015	-0.028***	0.017**	-0.025**	0.019*	-0.023**	0.018*	-0.024**	0.018*	-0.024**	0.022**	-0.023**	0.022**	-0.021*	0.021**	-0.021*
	(0.011)	*(0.012)*	*(0.011)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.009)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*
Years of Crime	-0.007***	0.026***	-0.007***	0.022***	-0.005**	0.020***	-0.005*	0.020***	-0.007***	0.022***			-0.004	0.016**					-0.012***	0.021***	-0.012***	0.020***
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.002)*	*(0.003)*			*(0.008)*	*(0.008)*					*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Years of Education	0.007	0.001	0.023***	0.001	-0.007*	0.004	-0.002	0.001					0.007	0.001	0.006	0.001	0.013***	-0.003	0.006	0.003	0.006	0.003
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.005)*					*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Schools per Young Person	0.331***		0.183**		0.324***		0.325***		0.127**		0.327***		0.332***		0.336***		0.280***		0.287***		0.295***	
	(0.071)		*(0.072)*		*(0.067)*		*(0.067)*		*(0.063)*		*(0.071)*		*(0.071)*		*(0.071)*		*(0.070)*		*(0.070)*		*(0.070)*	
Lagged Enrollment	0.189***		0.173***		0.199***		0.224***						0.189***		0.189***		0.169***		0.180***		0.182***	
	(0.012)		*(0.013)*		*(0.011)*		*(0.013)*						*(0.012)*		*(0.012)*		*(0.013)*		*(0.013)*		*(0.013)*	
Enrollment		0.099**		0.071***										0.103**		0.109**		0.139***				
		(0.050)		*(0.016)*										*(0.048)*		*(0.048)*		*(0.054)*				
Lagged Crime		0.156***		0.147***		0.144***		0.159***		0.141***				0.141***		0.140***		0.110***		0.141***		0.142***
		(0.013)		*(0.012)*		*(0.011)*		*(0.014)*		*(0.012)*				*(0.012)*		*(0.012)*		*(0.014)*		*(0.012)*		*(0.012)*
Cognitive Factor	0.043*	0.061**	0.038*	0.019	0.008	0.029	0.005	0.030	0.006	0.035	0.042*	0.033	0.043*	0.030	0.042*	0.029	0.041*	0.029	0.036	0.034	0.037	0.033
	(0.023)	*(0.024)*	*(0.022)*	*(0.022)*	*(0.021)*	*(0.022)*	*(0.021)*	*(0.022)*	*(0.018)*	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*	*(0.022)*	*(0.024)*	*(0.023)*	*(0.024)*
Social/Emotional Factor	0.008	-0.106***	0.020	-0.071***	-0.011	-0.068***	-0.012	-0.067***	0.008	-0.075***	0.006	-0.074***	0.007	-0.074***	0.007	-0.073***	0.012	-0.073***	0.024*	-0.070***	0.023	-0.071***
	(0.014)	*(0.015)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.011)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*
Jail					0.100***	0.091***	0.375***	0.008														
					(0.012)	*(0.013)*	*(0.077)*	*(0.089)*														
Enrollment (alternative)						0.056		0.048														
						(0.047)		*(0.047)*														

Table O5: Average Marginal Effects from Probits for Crime and Education (Violent Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Crime * Jail							0.000	-0.001														
							(0.005)	*(0.005)*														
Years of Education * Jail							-0.020***	0.009														
							(0.006)	*(0.007)*														
Lagged Enrollment * Jail							-0.088***															
							(0.023)															
Lagged Crime * Jail								-0.049**														
								(0.025)														
Enrollment * Jail								0.019														
								(0.026)														
Years of Education (alternative)									0.015***	0.006												
									(0.004)	*(0.004)*												
Lagged Enrollment (alternative)									0.085***													
									(0.010)													
Enrolment (alternative)										0.128**												
										(0.062)												
Years of Crime * Age1											-0.007**	0.020***										
											(0.003)	*(0.004)*										
Years of Crime * Age2											-0.007**	0.022***										
											(0.003)	*(0.003)*										
Years of Education * Age1											0.002	0.004										
											(0.005)	*(0.005)*										
Years of Education * Age2											0.009**	0.002										
											(0.004)	*(0.005)*										
Lagged Enrollment * Age1											0.229***											
											(0.020)											
Lagged Enrollment * Age2											0.172***											
											(0.016)											
Enrollment * Age1												0.080										
												(0.050)										
Enrollment * Age2												0.044										
												(0.058)										
Lagged Crime * Age1												0.140***										
												(0.018)										
Lagged Crime * Age2												0.146***										
												(0.018)										
Years of Crime Squared													-0.000	0.000								
													(0.001)	*(0.001)*								
Years of Crime: 0 to 4															-0.015***	0.031***						
															(0.006)	*(0.007)*						
Years of Crime: 5 to 9															-0.011***	0.023***						
															(0.003)	*(0.004)*						
Years of Crime: 10 or more															-0.005*	0.025***						
															(0.003)	*(0.003)*						
Years of Crime * Age of Entry 14																	-0.048***	0.031***				
																	(0.007)	*(0.007)*				

Table O5: Average Marginal Effects from Probits for Crime and Education (Violent Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Crime * Age of Entry 15																	-0.024***	0.048***				
																	(0.006)	*(0.006)*				
Years of Crime * Age of Entry 16																	-0.020***	0.044***				
																	(0.006)	*(0.006)*				
Years of Crime * Age of Entry 17																	0.009	0.056***				
																	(0.006)	*(0.008)*				
Years of Crime * Age of Entry 18																	0.006	0.071***				
																	(0.013)	*(0.013)*				
Crime																			0.202***		0.188***	
																			(0.051)		*(0.054)*	
Lagged State Arrest Rate																				-1.550		
																				(0.943)		
Rho	-0.141		-0.083***		-0.040		-0.037		-0.273**		-0.070		-0.155		-0.169		-0.237**		-0.402***		-0.442***	
	(0.103)		*(0.032)*		*(0.098)*		*(0.099)*		*(0.138)*		*(0.109)*		*(0.104)*		*(0.104)*		*(0.120)*		*(0.146)*		*(0.141)*	
Observations	5,232	5,232	6,236	6,236	6,236	6,236	6,236	6,236	5,139	5,139	5,232	5,232	5,232	5,232	5,232	5,232	5232	5232	5232	5232	5232	5232

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
2. In column (1) we do not include drug use as an independent regressor. In column (2), enrollment is set to zero if an individual did not attend a community school. In column (3), we condition on whether the individual is interviewed in jail, and in column (4) we interact the jail dummy with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail. In column (5) enrollment is redefined as attending school for at least nine months. Coefficients are allowed to vary by age in specification (6). Age1 is a dummy for ages 14 to 19, and Age2 is a dummy for ages 20 and above. In column (7) we use a quadratic function in criminal experience. In column (8) we use a piecewise-linear function of criminal experience: 0 to 4, 5 to 9, more than 10. In column (9) we use the criminal experience observed in the sample only, interacted with age of entry dummies. In column (10) we change the direction of the contemporaneous effect; we estimate the contemporaneous effect of crime on education. In column (11) we add the lagged state arrest rate as an exclusion in the crime equation.

Table O6: Average Marginal Effects from Probits for Crime and Education (Property Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.055**	0.059***	0.076***	0.076***	0.061***	0.076***	0.060***	0.076***	0.018	0.060***	0.057***	0.064***	0.054**	0.062***	0.053**	0.063***	0.064***	0.072***	0.049**	0.050	0.049**	0.059***
	(0.021)	*(0.018)*	*(0.021)*	*(0.016)*	*(0.020)*	*(0.016)*	*(0.020)*	*(0.016)*	*(0.018)*	*(0.017)*	*(0.021)*	*(0.017)*	*(0.021)*	*(0.017)*	*(0.021)*	*(0.017)*	*(0.021)*	*(0.018)*	*(0.022)*	*(0.054)*	*(0.022)*	*(0.016)*
Hispanic	-0.023	-0.014	-0.034**	-0.024**	-0.005	-0.026**	-0.005	-0.026**	-0.021*	-0.014	-0.023	-0.012	-0.023	-0.013	-0.023	-0.013	-0.026*	-0.020	-0.023	-0.011	-0.023	-0.011
	(0.015)	*(0.013)*	*(0.015)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.012)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Black	0.024	-0.010	-0.002	-0.017	0.043***	-0.019	0.040**	-0.019	0.002	-0.004	0.024	-0.005	0.025	-0.006	0.025	-0.005	0.020	-0.013	0.024	-0.006	0.024	-0.006
	(0.017)	*(0.016)*	*(0.017)*	*(0.014)*	*(0.016)*	*(0.014)*	*(0.016)*	*(0.014)*	*(0.013)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*
Other	0.036	0.008	0.010	-0.010	0.034	-0.008	0.031	-0.009	-0.024	0.012	0.036	0.016	0.037	0.014	0.036	0.016	0.014	0.012	0.034	0.014	0.034	0.014
	(0.027)	*(0.025)*	*(0.027)*	*(0.024)*	*(0.026)*	*(0.024)*	*(0.026)*	*(0.024)*	*(0.022)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*
Female	0.067***	-0.027*	0.126***	-0.023*	0.024*	-0.012	0.025*	-0.011	0.012	-0.026*	0.068***	-0.022	0.067***	-0.025*	0.067***	-0.025*	0.068***	-0.041***	0.067***	-0.027*	0.067***	-0.027*
	(0.014)	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*
Non-intact Family	-0.048***	0.007	-0.056***	0.003	-0.028**	0.000	-0.028**	-0.001	-0.003	0.009	-0.049***	0.007	-0.048***	0.007	-0.048***	0.007	-0.041***	0.008	-0.049***	0.008	-0.049***	0.008
	(0.015)	*(0.014)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.012)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Siblings	-0.003	0.001	-0.003	0.002	-0.002	0.002	-0.002	0.002	-0.004**	0.002	-0.003	0.002	-0.003	0.002	-0.003	0.002	-0.003	0.002	-0.003	0.002	-0.003	0.002
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.082***	-0.030***	-0.072***	-0.026***	-0.075***	-0.029***	-0.075***	-0.029***	-0.047***	-0.027***	-0.090***	-0.026***	-0.083***	-0.027***	-0.082***	-0.027***	-0.089***	-0.030***	-0.080***	-0.024***	-0.080***	-0.024***
	(0.004)	*(0.006)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.003)*	*(0.004)*	*(0.005)*	*(0.007)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.007)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.003)*
Certainty of Punishment	0.004	-0.017***	0.007***	-0.013***	-0.000	-0.013***	-0.000	-0.013***	0.006***	-0.015***	0.003	-0.015***	0.004	-0.016***	0.004	-0.016***	0.003	-0.015***	0.004*	-0.016***	0.004*	-0.016***
	(0.003)	*(0.002)*	*(0.003)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Children	-0.017**	-0.004	-0.016**	0.002	-0.016**	0.003	-0.014**	0.002	-0.023***	-0.001	-0.017**	0.000	-0.018**	0.001	-0.017**	0.001	-0.020***	0.001	-0.018**	0.001	-0.018**	0.001
	(0.007)	*(0.007)*	*(0.008)*	*(0.006)*	*(0.007)*	*(0.006)*	*(0.007)*	*(0.006)*	*(0.008)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Family Crime	0.000	0.112***	-0.035**	0.090***	0.008	0.085***	0.007	0.087***	0.010	0.096***	0.002	0.095***	0.002	0.095***	0.002	0.094***	0.001	0.097***	-0.008	0.094***	-0.007	0.094***
	(0.015)	*(0.011)*	*(0.014)*	*(0.010)*	*(0.014)*	*(0.010)*	*(0.014)*	*(0.010)*	*(0.011)*	*(0.011)*	*(0.015)*	*(0.011)*	*(0.015)*	*(0.011)*	*(0.015)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.017)*	*(0.011)*	*(0.017)*	*(0.011)*
Drug Use			0.040***	0.154***	-0.044***	0.155***	-0.045***	0.154***	-0.043***	0.139***	-0.006	0.143***	-0.007	0.144***	-0.006	0.143***	-0.005	0.145***	-0.016	0.144***	-0.016	0.144***
			(0.011)	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.009)*	*(0.010)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.014)*	*(0.009)*	*(0.014)*	*(0.009)*
Unemployment Rate	0.021***	0.010**	0.020***	0.012***	0.023***	0.013***	0.024***	0.013***	0.017***	0.010**	0.022***	0.011**	0.021***	0.010**	0.021***	0.010**	0.022***	0.010**	0.021***	0.009*	0.021***	0.009**
	(0.006)	*(0.005)*	*(0.006)*	*(0.004)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.004)*
Future Outlook Inventory	0.018	-0.040***	0.018*	-0.037***	0.014	-0.036***	0.014	-0.036***	0.016*	-0.033***	0.019*	-0.033***	0.018	-0.032***	0.017	-0.032***	0.020*	-0.030***	0.019*	-0.033***	0.019*	-0.033***
	(0.011)	*(0.010)*	*(0.011)*	*(0.009)*	*(0.010)*	*(0.009)*	*(0.010)*	*(0.009)*	*(0.009)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*
Years of Crime	-0.004	0.022***	-0.004*	0.018***	-0.002	0.018***	-0.002	0.019***	-0.004*	0.018***			-0.011*	0.025***					-0.005*	0.018***	-0.005*	0.018***
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.003)*	*(0.003)*	*(0.002)*	*(0.002)*			*(0.006)*	*(0.006)*					*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Years of Education	0.007*	0.007*	0.023***	0.006*	-0.007*	0.007**	-0.002	0.007*					0.007*	0.006	0.007*	0.006	0.012***	0.005	0.007	0.005	0.007	0.005
	(0.004)	*(0.004)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.004)*					*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Schools per Young Person	0.326***		0.177**		0.322***		0.322***		0.132**		0.328***		0.330***		0.329***		0.276***		0.323***		0.323***	
	(0.071)		*(0.072)*		*(0.067)*		*(0.067)*		*(0.064)*		*(0.072)*		*(0.071)*		*(0.071)*		*(0.071)*		*(0.071)*		*(0.071)*	
Lagged Enrollment	0.191***		0.175***		0.199***		0.223***						0.190***		0.190***		0.175***		0.191***		0.191***	
	(0.012)		*(0.013)*		*(0.011)*		*(0.013)*						*(0.012)*		*(0.012)*		*(0.012)*		*(0.012)*		*(0.012)*	
Enrollment		-0.029		-0.007										-0.022		-0.020		-0.002				
		(0.040)		*(0.013)*										*(0.040)*		*(0.040)*		*(0.044)*				
Lagged Crime		0.163***		0.142***		0.140***		0.156***		0.145***				0.144***		0.144***		0.126***		0.144***		0.144***
		(0.010)		*(0.010)*		*(0.010)*		*(0.012)*		*(0.010)*				*(0.010)*		*(0.010)*		*(0.011)*		*(0.010)*		*(0.010)*
Cognitive Factor	0.040*	0.049**	0.035	0.007	0.006	0.012	0.003	0.013	0.004	0.014	0.040*	0.022	0.040*	0.021	0.039*	0.022	0.037	0.030	0.039*	0.018	0.039*	0.018
	(0.023)	*(0.020)*	*(0.022)*	*(0.019)*	*(0.021)*	*(0.019)*	*(0.021)*	*(0.019)*	*(0.018)*	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*
Social/Emotional Factor	0.008	-0.092***	0.017	-0.065***	-0.011	-0.064***	-0.012	-0.064***	0.009	-0.065***	0.005	-0.068***	0.007	-0.068***	0.007	-0.068***	0.003	-0.076***	0.012	-0.067***	0.012	-0.067***
	(0.014)	*(0.013)*	*(0.014)*	*(0.012)*	*(0.013)*	*(0.012)*	*(0.013)*	*(0.012)*	*(0.011)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.012)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Jail					0.100***	0.045***	0.388***	0.061														
					(0.012)	*(0.011)*	*(0.077)*	*(0.074)*														
Enrollment (alternative)						-0.024		-0.035														
						(0.038)		*(0.038)*														

Table O6: Average Marginal Effects from Probits for Crime and Education (Property Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Crime * Jail							-0.002	-0.004														
							(0.004)	*(0.004)*														
Years of Education * Jail							-0.020***	0.002														
							(0.006)	*(0.006)*														
Lagged Enrollment * Jail							-0.086***															
							(0.023)															
Lagged Crime * Jail								-0.050**														
								(0.020)														
Enrollment * Jail								0.019														
								(0.021)														
Years of Education (alternative)									0.015***	0.005												
									(0.004)	*(0.004)*												
Lagged Enrollment (alternative)									0.090***													
									(0.010)													
Enrolment (alternative)										-0.074												
										(0.049)												
Years of Crime * Age1											-0.004	0.024***										
											(0.003)	*(0.003)*										
Years of Crime * Age2											-0.003	0.014***										
											(0.003)	*(0.003)*										
Years of Education * Age1											0.003	0.005										
											(0.005)	*(0.004)*										
Years of Education * Age2											0.010**	0.007*										
											(0.004)	*(0.004)*										
Lagged Enrollment * Age1											0.229***											
											(0.020)											
Lagged Enrollment * Age2											0.172***											
											(0.016)											
Enrollment * Age1												-0.035										
												(0.041)										
Enrollment * Age2												-0.050										
												(0.046)										
Lagged Crime * Age1												0.130***										
												(0.014)										
Lagged Crime * Age2												0.157***										
												(0.016)										
Years of Crime Squared													0.001	-0.001								
													(0.001)	*(0.001)*								
Years of Crime: 0 to 4															-0.008*	0.024***						
															(0.004)	*(0.005)*						
Years of Crime: 5 to 9															-0.004*	0.020***						
															(0.002)	*(0.003)*						
Years of Crime: 10 or more															0.000	0.016***						
															(0.005)	*(0.003)*						
Years of Crime * Age of Entry 14																	-0.055***	0.020***				
																	(0.009)	*(0.007)*				

Table O6: Average Marginal Effects from Probits for Crime and Education (Property Crime) - Robustness Checks

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance		Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education -- Not Instrumenting		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Crime * Age of Entry 15																	-0.033***	0.039***				
																	(0.007)	*(0.006)*				
Years of Crime * Age of Entry 16																	-0.022***	0.036***				
																	(0.007)	*(0.006)*				
Years of Crime * Age of Entry 17																	0.008	0.040***				
																	(0.007)	*(0.006)*				
Years of Crime * Age of Entry 18																	0.021	0.050***				
																	(0.016)	*(0.012)*				
Crime																			0.058		0.057	
																			(0.051)		*(0.051)*	
Lagged State Arrest Rate																				-0.137		
																				(0.829)		
Rho	0.067		0.006		0.044		0.055		0.172		0.122		0.061		0.055		0.012		-0.130		-0.133	
	(0.110)		*(0.036)*		*(0.107)*		*(0.107)*		*(0.147)*		*(0.121)*		*(0.118)*		*(0.118)*		*(0.127)*		*(0.121)*		*(0.122)*	
Observations	5,232	5,232	6,231	6,231	6,231	6,231	6,231	6,231	5,141	5,141	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5232	5232	5232	5232

Notes:

1. Standard errors are reported below the point estimates in italics and in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.

2. In column (1) we do not include drug use as an independent regressor. In column (2), enrollment is set to zero if an individual did not attend a community school. In column (3), we condition on whether the individual is interviewed in jail, and in column (4) we interact the jail dummy with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail. In column (5) enrollment is redefined as attending school for at least nine months. Coefficients are allowed to vary by age in specification (6). Age1 is a dummy for ages 14 to 19, and Age2 is a dummy for ages 20 and above. In column (7) we use a quadratic function in criminal experience. In column (8) we use a piecewise-linear function of criminal experience: 0 to 4, 5 to 9, more than 10. In column (9) we use the criminal experience observed in the sample only, interacted with age of entry dummies. In column (10) we change the direction of the contemporaneous effect; we estimate the contemporaneous effect of crime on education. In column (11) we add the lagged state arrest rate as an exclusion in the crime equation.