

Mancino, Maria Antonella; Navarro, Salvador; Rivers, David A.

Working Paper

Separating state dependence, experience, and heterogeneity in a model of youth crime and education

CIBC Working Paper, No. 2015-1

Provided in Cooperation with:

Centre for Human Capital & Productivity (CHCP), Department of Economics, University of Western Ontario

Suggested Citation: Mancino, Maria Antonella; Navarro, Salvador; Rivers, David A. (2015) : Separating state dependence, experience, and heterogeneity in a model of youth crime and education, CIBC Working Paper, No. 2015-1, The University of Western Ontario, CIBC Centre for Human Capital and Productivity, London (Ontario)

This Version is available at:

<https://hdl.handle.net/10419/121968>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

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

by

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

Working Paper # 2015-1

March 2015



CIBC Working Paper Series

Department of Economics
Social Science Centre
The University of Western Ontario
London, Ontario, N6A 5C2
Canada

This working paper is available as a downloadable pdf file on our website
<http://www.ssc.uwo.ca/economics/centres/cibc/>

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*

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 that feeds 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 for crime and schooling. We provide evidence that, as a consequence, policies that affect individual heterogeneity (like 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

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

Understanding what determines crime among youths can be very useful for designing policies to reduce not only youth crime, but crime rates overall. The empirical evidence suggests that youth account for a large share of crime in the United States. In 1997, law enforcement officials arrested 2.8 million people under the age of 18, accounting for one fifth of all arrests (Snyder and Sickmund, 1999). There is also evidence that crime is becoming increasingly concentrated among teenagers. The rate at which juveniles were arrested for murder and violent crime rose 177% and 79%, respectively between 1978 and 1993, about three times the increase for adults over the same period (Levitt, 1998). Furthermore, numerous studies have found that criminal activity is highly persistent over time ((Blumstein, Farrington, and Moitra, 1985; Nagin and Paternoster, 2000, 1991)), which suggests that attempts to reduce youth crime may have large effects on adult crime as well.¹

In this paper we study the determinants of youth crime in the context of a joint dynamic discrete choice model of crime and education in which previous and current education and criminal activities affect current crime and educational decisions. Understanding the relationships between crime and education has important policy implications. To the extent that education and crime interact, this provides additional instruments for policy makers interested in reducing crime and/or increasing educational attainment.

Recently there has been an increased recognition in the literature that education may be an important driver of criminal behavior, in a variety of ways. Increased educational attainment may increase future wages, which increases the return to legitimate work and raises the opportunity cost of illicit behavior (Freeman, 1996; Lochner, 2004). Schooling may alter people's preferences, for example by increasing patience or risk aversion (Becker and Mulligan, 1997; Usher, 1997; Fajnzylber, Lederman, and Loayza, 2002). Education may directly affect the psychic or financial rewards from crime by emphasizing social and emotional development (Lochner, 2011). Schooling can also have an incapacitation effect (Freeman, 1996; 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.² Schooling can affect social networks, and these networks could influence criminal behavior, for example via gang

¹In 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 (Jacob and Lefgren, 2003; Grogger, 1998, 1995). Moreover, higher levels of crime have been found to reduce incentives for investment (Zeilekha and Bar-Efrat, 2009).

²The 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 schooling is negatively associated with both property and violent crime rates.

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). This could also, in turn, feed back into crime choices.

Because we observe individuals over many years, we are able to account for both of these possibilities. 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 for different crime categories as well as his educational experience and enrollment decisions over time. This type of data is ideally suited for understanding the dynamics in crime and education since it follows individuals over time. The enrolled youth were between 14 and 18 years old at the time of their committing offense and were found guilty of a serious offense. Each study participant was followed for a period of seven years past enrollment which results in a comprehensive picture of life changes in a wide array of areas over the course of this time.³

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 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 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, 2009), the Philadelphia Birth Cohort Study (Imai and Krishna, 2004; Tauchen, Witte, and Griesinger, 1994), the National Youth Survey (Imai, Katayama, and Krishna, 2006), the National Longitudinal Study of Adolescent Health (Mocan and Rees, 2005), among others. A common feature of

³We describe the dataset in more detail in Section 2.

these datasets is that they study subsets of the population at large, and are not focused particularly on criminals.⁴ For instance, the participation rate in violent and property crime in the NLSY79 averages 6.8% and 11.8% for individuals aged between 20 and 23 years old with 12 years of schooling (Lochner, 2004).

While these datasets are useful for examining what drives people over the threshold between non-criminal and criminal behavior, they are not well-suited for studying serious criminal behavior, as they include very few serious offenders and few serious crimes. For policy makers interested in reducing overall crime rates, particularly violent crime rates, then data on these serious offenders, who contribute significantly to aggregate crime rates, is necessary.

An advantage of studying only serious offenders through the PD data, is that to the extent there is unobserved heterogeneity that leads some individuals to become serious offenders, we are more likely to be working with people who are on a criminal trajectory (adolescence-limited, high-level chronics and low-level chronics as defined by Nagin and Land, 1993 and Nagin, Farrington, and Moffitt, 1995). 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 done so.

The PD data includes a much richer set of targeted control variables than is typically available. In addition to standard socio-economic variables and information about their 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 committed 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.

An additional benefit of this dataset is that, in addition to very detailed survey questions, individuals are also 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 ability (sometimes referred to as "non-cognitive ability") on outcomes such as labor market participation, test scores, health, and migration (Borghans, Meijers, and Ter Weel, 2008; Heckman, Stixrud, and Urzua, 2006; Cobb-Clark and Tan, 2011; Chiteji,

⁴An exception is the dataset used by Witte (1980) which follows a set of men released from prison in North Carolina. That dataset covers only adults and does not incorporate education decisions.

⁵Empirical 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).

2010; Jaeger et al., 2010), which suggests a possible role for social/emotional skills in predicting criminal activity as well. Another strand of empirical research has focused on understanding the factors that affect the evolution of cognitive and non-cognitive skills and how relevant this is for building optimal interventions (Cunha et al., 2006; Cunha and Heckman, 2007; Cunha, Heckman, and Schennach, 2010).

Focusing specifically on crime, Hill et al. (2011) show that several interventions that focus on personality rather than cognitive skills were effective at reducing delinquency and traits related to delinquency. Also, Heckman, Stixrud, and Urzua (2006) show that both cognitive and non-cognitive skills influence risky activities such as smoking by age 18, imprisonment, and participation in illegal activities. Although the economic literature on the effect of social/emotional skills on crime is limited, research from criminology and psychology has estimated correlations between IQ, measures of personality, and crime. This literature finds significant correlations between personality measures such as self-control, restraint, negative emotionality and crime/delinquency (Caspi et al., 1994; Agnew et al., 2002; Pratt and Cullen, 2000).

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 commonly observed in datasets, such as drug use, involvement in crime by family members, attitudes towards the future, and both cognitive 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 also evidence of returns to experience as well.

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 discuss the policy implications of our results and provide some policy simulations. 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) in the juvenile or adult court systems in Maricopa County, Arizona, or Philadelphia County, Pennsylvania. The study follows 1,354 individuals, who were at least 14 years old and under 18 years old at the time of their offense. Besides fulfilling the requirements of age and criminal activity, individuals had to provide informed assent or consent to participate in the study.⁷ Individuals were paid \$50 to participate in the baseline survey, with compensation increasing for the follow-ups to minimize attrition (Monahan et al., 2009). Twenty percent of the youths approached for participation declined.

The initial (baseline) survey occurred when individuals first entered the sample. There were six semi-annual follow-up interviews, followed by four annual follow-up interviews. In total, the survey follows each individual for a period of eight years.

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 rich 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 (intact family),⁸ and whether any family members are involved in criminal activities.⁹

We also observe whether individuals use drugs, as well as their perceived risk to offending (the individual-specific perceived probability of being caught).¹⁰ Furthermore, we have a measure of how much individuals care about the future, though a variable called the *Future Outlook Inventory*. This

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

⁷Parental consent was obtained for all youth younger than 18 at the time of enrollment in the survey.

⁸

Dornbush et al. (1985) have shown that family composition during childhood may affect delinquent behavior.

⁹

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).

¹⁰

Each individual's perceived risk is measured in each period by asking them how likely it is that they will be caught and arrested for the following seven crimes: fighting, robbery with 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). We use the average of these seven responses in our empirical analysis.

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 complement this with data from the Bureau of Labor Statistics on local annual unemployment rates, and with annual data on the number of schools per person from the National Center of Education Statistics, and on the number of people between the ages of 15 and 19 in each county from the U.S. Census.¹¹

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 (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.¹²

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 subsets: 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 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.¹³

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

¹¹We use the latter two to compute the number of schools per person of high school age in each county-year pair.

¹²The cognitive tests are given only in the baseline survey, whereas the social/emotional tests are repeated in the follow-up surveys as well.

¹³Higher scores indicate disability or impairment.

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 year. 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 are triggered that 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 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 data restrictions, and therefore are not included in our analysis.¹⁵ We also compute results 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 do not know the criminal history prior to the initial survey, with the exception of knowing the age at which each individual first committed each of the crimes. Since the total years of accumulated crime

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

¹⁵Not 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.

will be an important variable in our model, this generates a missing data problem. For individuals who first committed a crime after the initial survey, we can accurately compute the total years of criminal experience in each year, since we can infer their entire history of choices. However, some people started criminal activities as early as age 9. 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 at the age of the baseline survey. Experience in subsequent years is then calculated based on this estimate and on the observed crime decisions.¹⁶

We construct four panel datasets, one for each of the three crime measures described above and one with all crime combined together. Table 1 reports descriptive statistics for our four samples. There are a couple of statistics that we want to highlight. First, crime rates in the sample are quite high. The violent crime rate is 43.5%, 28.5% for property crime, and 21.3% 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, and there is a large percentage of minorities, with blacks and Hispanics representing 40% and 34%, respectively. Drug use is also quite prominent, with an average of 47%. Another interesting observation is that 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 and social/emotional 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 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, 20% have some level of cognitive impairment according to Trail-Making A, and 38% under

¹⁶An alternative to this imputation procedure is to, at estimation time, use the probabilities predicted by our model to integrate the likelihood for each individual starting at age 9 (the earliest age at which any individual started criminal activities) up until the age of the first interview. This substantially increases the computational burden. We tried this for our baseline specification described in Section 4 below, and found that the results were very similar.

Trail-Making B. The Stroop Tests take a continuum of values, and 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.

The raw social/emotional test scores, described in Table 3, are a little more difficult 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). This suggests that the WAI scores may be more likely to be related to criminal behavior, which is consistent with our empirical analysis below.

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. Figure 2 shows the same for education. There are two important relationships to notice. 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 even 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 ability youths 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 had learned the previous year’s material. A third possibility is that there are returns to previous experience. For example, 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, particularly in the mid to late teenage years. While this would seem to suggest a negative effect of 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. The utility associated with the choice $e_{i,t}$ is a function of all relevant decision variables including years of crime and years of education up to t ($YC_{i,t}$ and $YE_{i,t}$), and a set of individual-specific characteristics ($Z_{i,t}$):

$$U_{i,t}^e(e_{i,t}) = Z_{i,t}\beta_t^e e_{i,t} + YC_{i,t}\lambda_t^e e_{i,t} + YE_{i,t}\alpha_t^e + \xi_{i,t}^e(e_{i,t}), \quad (1)$$

where $\xi_{i,t}^e(c_{i,t})$ denotes unobservable individual-specific utility terms. The net utility of getting education in period t is:

$$V_{i,t}^e := U_{i,t}^e(e_{i,t} = 1) - U_{i,t}^e(e_{i,t} = 0) = Z_{i,t}\beta_t^e + YC_{i,t}\lambda_t^e + YE_{i,t}\alpha_t^e + \eta_{i,t}^e, \quad (2)$$

where $\eta_{i,t}^e = \xi_{i,t}^e(1) - \xi_{i,t}^e(0)$. An individual chooses to enroll in school ($e_{i,t} = 1$) if and only if $V_{i,t}^e > 0$.

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

$$U_{i,t}^c(c_{i,t}) = Z_{i,t}\beta_t^c c_{i,t} + YE_{i,t}\alpha_t^c c_{i,t} + E_{i,t}\gamma_t^c c_{i,t} + YC_{i,t}\lambda_t^c c_{i,t} + \xi_{i,t}^c(c_{i,t}) \quad (3)$$

The net expected utility of crime commission is:

$$V_{i,t}^c := U_{i,t}^c(c_{i,t} = 1) - U_{i,t}^c(c_{i,t} = 0) = Z_{i,t}\beta_t^c + YE_{i,t}\alpha_t^c + E_{i,t}\gamma_t^c + YC_{i,t}\lambda_t^c + \eta_{i,t}^c \quad (4)$$

where $\eta_{i,t}^c = \xi_{i,t}^c(1) - \xi_{i,t}^c(0)$. Given the enrollment decision, the individual chooses to commit crime

$(c_{i,t} = 1)$ if and only if $V_{i,t}^c > 0$.

Notice that in equations (2) and (4) 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.¹⁷ 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.¹⁸

Imposing the coherency condition makes our model triangular, which allows us to factor the likelihood in the following way:¹⁹

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

where $\Pr(C_{i,t} = 1 \mid E_{i,t} = e_{i,t}) = \Pr(V_{i,t}^c > 0 \mid E_{i,t} = e_{i,t})$ and $\Pr(E_{i,t} = 1) = \Pr(V_{i,t}^e > 0)$, 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 (2) and (4) 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 three strategies. First, we include the number of schools per student (by county and year), as a measure of the cost of attending school, 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 for identification of the effect of enrollment on crime.

Second, 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

¹⁷See Heckman (1978) and Tamer (2003) for further discussion of the identification problems associated with dummy endogenous variables in simultaneous equations models

¹⁸Alternatively we could assume that the contemporaneous effect of enrollment on crime is zero. We chose this assumption because the literature is focused more on the effect of education on crime, as opposed to the effect of crime on education.

¹⁹We keep the conditioning on the remaining variables implicit to ease on notation.

equations (2) and (4) as follows:

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

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

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 out of the error term components 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 $C_{i,t}$ and $E_{i,t}$ where

$$\begin{aligned} C_{i,t} &= \begin{cases} 1 & \text{if } V_{i,t}^c > 0 \\ 0 & \text{otherwise} \end{cases}, \\ E_{i,t} &= \begin{cases} 1 & \text{if } V_{i,t}^e > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

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

$$\begin{aligned} V_{i,t}^c &= Z_{i,t} \beta_t^c + Y E_{i,t} \alpha_t^c + E_{i,t} \gamma_t^c + Y C_{i,t} \lambda_t^c + \delta^{c,cog} \bar{\theta}_i^{cog} + \delta^{c,emo} \bar{\theta}_i^{emo} + \varepsilon_{i,t}^c, \\ V_{i,t}^e &= Z_{i,t} \beta_t^e + Y C_{i,t} \lambda_t^e + Y E_{i,t} \alpha_t^e + W_t + \delta^{e,cog} \bar{\theta}_i^{cog} + \delta^{e,emo} \bar{\theta}_i^{emo} + \varepsilon_{i,t}^e, \end{aligned}$$

and where

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

3.1 Factor Model for Abilities

Let $M_{j,i}^{cog}$ denote one of $j = 1, \dots, J$ cognitive measurements. In particular, 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, 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 $k = 1, \dots, K$ continuous tests of social/emotional skills that are repeatedly measured over time, which we denote by $M_{k,i,t}^{beh}$. 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

$$\begin{aligned} M_{j,i}^{cog} &= Z_{i,1}\beta_j^{cog} + \theta_i^{cog}\delta_j^{cog} + \xi_{j,i}^{cog}, \\ M_{k,i,t}^{emo} &= Z_{i,t}\beta_{k,t}^{emo} + \theta_i^{emo}\delta_{k,t}^{emo} + \xi_{k,i,t}^{emo}, \end{aligned}$$

and we write an ordered model

$$M_{j,i}^{cog} = \mathbb{1}(\psi_l < Z_{i,1}\beta_j + \theta_i^{cog}\delta_j^{cog} + \xi_{j,i}^{cog} \leq \psi_{l+1})$$

for the discrete Trail-Making measures. $\theta_i^{cog}, \theta_i^{emo}$ denote cognitive and social/emotional abilities respectively, $\delta_j^{cog}, \delta_{k,t}^{emo}$ denote loadings that measure the effect of these skills, and the “uniquenesses” $\{\xi_{j,i}^{cog}\}_{j=1}^J, \left\{\left\{\xi_{k,i,t}^{emo}\right\}_{t=1}^T\right\}_{k=1}^K$ capture other determinants of the test scores like measurement error. While we assume that θ_i^{cog} and θ_i^{emo} 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).²⁰ It begins by (conditionally) demeaning the measurements (i.e., recovering the β 's). It then shows that the loadings (i.e., the δ 's) are identified by taking covariances between different cognitive measures and between different social/emotional measures. The marginal distributions of θ_i^{cog} and $\{\xi_{j,i}^{cog}\}_{j=1}^J$, as well as those of θ_i^{emo} and $\left\{\left\{\xi_{k,i,t}^{emo}\right\}_{t=1}^T\right\}_{k=1}^K$ 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.

However, for estimation purposes, we impose distributional assumptions. In particular, we assume that

²⁰Given that the θ_i 's are latent, we need to normalize one δ_i for each type of skills to some number, e.g., $\delta_{1,1}^{cog} = \delta_{1,1}^{emo} = 1$. This normalization implies no restriction since a) we need to pin down the sign of δ_1^{cog} for interpretation as having more of θ_i^{cog} with a $\theta_i^{cog}\delta_1^{cog} > 0$ is equivalent to having less with $\theta_i^{cog}\delta_1^{cog} < 0$, and b) we need to pin down the variance of θ_i^{cog} as $\theta_i^{cog}\delta_1^{cog} = \theta_i^{cog}\kappa\frac{\delta_1^{cog}}{\kappa}$ for any constant κ . The same holds true for θ_i^{emo} .

$\xi_{j,i}^{cog} \sim N(0, \sigma_{\xi,cog,j}^2)$, $\xi_{k,i,t}^{emo} \sim N(0, \sigma_{\xi,emo,k}^2)$ and that

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

Given these distributional assumptions, the factor model is estimated by maximum likelihood. Let $\mathcal{M}_{j,i}^{cog} = M_{j,i}^{cog} - Z_{i,1}\beta_j^{cog} - \theta_i^{cog}\delta_j^{cog}$ and $\mathcal{M}_{k,i,t}^{emo} = M_{k,i,t}^{emo} - Z_{i,t}\beta_{k,t}^{emo} - \theta_i^{emo}\delta_{k,t}^{emo}$. The contribution to the likelihood of observation i is given by

$$\int \int \left[\frac{\prod_{j=1}^J \phi(\mathcal{M}_{j,i}^{cog} | \theta_i^{cog}; \sigma_{\xi,cog,j}^2) \times \prod_{t=1}^T \prod_{k=1}^K \phi(\mathcal{M}_{k,i,t}^{emo} | \theta_i^{emo}; \sigma_{\xi,emo,k}^2)}{\prod_{t=1}^T \prod_{k=1}^K \phi(\mathcal{M}_{k,i,t}^{emo} | \theta_i^{emo}; \sigma_{\xi,emo,k}^2)} \right] \Phi(\theta_i^{cog}, \theta_i^{emo}; \sigma_{\theta,cog}^2, \sigma_{\theta,emo}^2, \rho) d\theta_i^{cog} d\theta_i^{emo},$$

where $\phi(|; \sigma^2)$ is the pdf of a mean zero normal with variance σ^2 and $\Phi(; a, b, c)$ is the pdf of a 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

$$\begin{aligned} \begin{pmatrix} \bar{\theta}_i^{cog} \\ \bar{\theta}_i^{emo} \end{pmatrix} &= \int \int \begin{pmatrix} \theta_i^{cog} \\ \theta_i^{emo} \end{pmatrix} f(\theta_i^{cog}, \theta_i^{emo} | M_i, Z_i; \hat{\beta}, \hat{\delta}, \hat{\sigma}, \hat{\rho}) d\theta_i^{cog} d\theta_i^{emo} \\ &= \int \int \begin{pmatrix} \theta_i^{cog} \\ \theta_i^{emo} \end{pmatrix} \frac{f(M_i | \theta_i^{cog}, \theta_i^{emo}, Z_i; \hat{\beta}, \hat{\delta}, \hat{\sigma}, \hat{\rho})}{f(M_i | Z_i; \hat{\beta}, \hat{\delta}, \hat{\sigma}, \hat{\rho})} d\theta_i^{cog} d\theta_i^{emo} \end{aligned}$$

which follows directly from our expression for the likelihood.

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 ability, and one related to social/emotional ability.

4.1 Factor Analysis

The results from the estimation of the factor model are presented in Tables 4-6 and Figure 4. We chose the following normalizations. The factor representing cognitive skills is normalized to have a loading of one in the matrix reasoning WASI test score, while for the factor representing social/emotional skills the loading is normalized to one on 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 1 standard deviation in cognitive skills leads to an increase of 1 standard deviation on the matrix reasoning WASI test, and similarly for the social/emotional factor.

In Table 4 we present the estimated distribution of the factors. 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 than in cognitive skills, and that they are positively correlated with a correlation coefficient around 0.23.

In Figure 4 we also 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 factor and the proportion contributed by the uniqueness.²¹ In Tables 5 and 6 we present the estimated parameters of the factor model for the measurement system.

As can be seen from Table 5, having more cognitive ability 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, our measure of cognitive skills is more related to the Stroop measures of cognitive dysfunction than to the WASI-IQ measures. However, even for the Stroop measures, cognitive skills can only explain at most 62% of the unobserved variance.

As documented in Table 6, 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 ability becomes 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.

²¹In order to avoid having a graph for each of the 8 periods for which we have social/emotional measures, we use the time-averaged loadings in our calculations.

What this result suggests is 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 in order 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 factor estimates as regressors.²² The results from the baseline bivariate probit are listed in column 1 of Table 7, 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.²³ The results for drug-related, violent, and property crime separately are found in Tables 8, 9, and 10, respectively.

We find that being in Maricopa County (compared to Philadelphia County) is associated with a higher probability of enrollment in school. The effect of location on crime varies by crime category. Violent and property crime are higher (2.6 and 5.2%-points respectively) and drug crime is lower (1.8%-points) in Maricopa County. 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 (10.2%-points).

Consistent with what one would expect, having a “non-intact” family, is associated with lower enrollment rates and higher crime rates, particularly for drug-related crimes. Age is negatively associated with enrollment, and both property and violent crime. The result for enrollment is not surprising given that this dataset covers people between the ages of 14 and 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). The fact that 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, 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).

Not surprisingly the effect of the perceived risk of punishment has no effect on education and has

²²As a robustness check, in Section 4.3.5 we use the set of measurements which we use to infer the factors, as regressors directly.

²³Note 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.

a negative effect on crime. The results are similar across crimes, 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.²⁴ Each child an individual has decreases the probability of enrollment by about 1.6%-points, but has no effect on crime. Having family members involved in crime has a large positive effect on crime (14.9%-points), with the largest effect for violent crime, 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 (0.2%-points). It has a large positive effect, however, on overall crime (22.4%-points). The effect on drug-related crime (21.3%-points) is the highest, but the effects on violent crime (15.9%-points) and property crime (14.4%-points) are also quite large.

We also include the unemployment rate to control for local employment conditions, and find that it has a positive effect on education: an increase in the unemployment rate by one percentage point leads to an increased probability of enrollment by 2.1%-points. The effect of unemployment on crime is also positive but small in magnitude and not statistically significant. These results suggest that criminal youth respond to worsening economic conditions by staying in school, rather than seeking employment opportunities, but this does not lead to further increases in criminal behavior. Our results are consistent with those of Betts and McFarland (1995) and Dellas and Sakellaris (2003) who estimate that a 1 percent increase in the unemployment rate increases enrollment in college by 2 to 4 percent. With regards to crime, a number of recent studies (e.g., Raphael and Winter-Ebmer, 2001 and Gould, Weinberg, and Mustard, 2002) find a sizable positive effect of unemployment on crime, although Lochner (2004) finds that the local unemployment rate has no effect on crime.

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; Akerlund et al., 2014), 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 are associated with more positive outcomes. Therefore, we should expect them to be positively associated with education and negatively associated with crime. We find that higher cognitive ability increases the

²⁴These 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 percent and in auto theft by more than 8 percent.

likelihood of enrollment and higher social/emotional skills lead to lower crime rates. The results imply that a one standard deviation increase in cognitive skills leads to an increased probability of enrollment of 1.3%-points, and a one standard deviation increase in social/emotional skills leads to a decrease in the probability of crime of 3.5%-points. The effects of cognitive ability 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 Vytlačil, 2001; Heckman, Stixrud, and Urzua, 2006; Murnane, Willett, and Levy, 1995, Herrnstein and Murray, 1994). 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 more compressed relative to the population at large. As a result, a one standard deviation change is not particularly large.

In addition to controlling for many source of individual heterogeneity, we also 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.²⁵ 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, 2009).

We find strong evidence of state dependence in both the education and crime decisions. Enrolling in school the previous period increases the probability of enrolling in the current period by 18.8%-points. Participating in crime in the previous period increases the probability of crime by 15.9%-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 2.1%-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.

²⁵For 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.

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 to intervene. We discuss this more in Section 5 when we illustrate these effects with various policy simulations based on our model.

We also examine the effect of contemporaneous education on crime. We find that the effect varies by the type of crime. For property crime, we find weak evidence that enrolling in school decreases crime, with an average marginal effect of 1.8%-points that is not precisely estimated. This is consistent with the incapacitation effect found by Jacob and Lefgren (2003) and Luallen (2006), although our effect is much smaller in magnitude. One explanation for this is that our dataset consists of serious offenders only. It is possible that being in school has stronger incapacitation effects for more minor property crimes, which are represented proportionally much more in datasets that sample from the population at large.

For violent and drug-related crime, we find the opposite effect: enrollment leads to an increase in crime rates (10.9%-points for violent and 7.9 for drug-related). This suggests the presence of 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.

In order to account for the possibility that education 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.

The last row of Table 7 reports the correlation in errors of the crime and enrollment equations. The estimate of -0.156 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.

4.3 Alternative Specifications

In this section we present results from a number of alternative specifications to our baseline model. The first set of results is designed to illustrate how our modeling choices affect the estimates. 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); do not allow for dynamics; and do not include the number of schools per student as an exclusion in the enrollment equation.

The objective for the second set of results is to provide some robustness checks to the baseline model.²⁶ We show that our results are robust to using the direct measures of cognitive and social/emotional skills, as opposed to our estimates of the underlying skills from the factor model; alternative ways to treat decisions while in jail; alternative definitions of enrollment; and allowing the effects of prior crime and education decisions, as well as contemporaneous enrollment, to vary by age.

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, 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 (county, gender, and age). The results are reported in column 2 of Table 7. 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 2.1 to 3.9%-points, and the effect of lagged crime increases by 50% from 15.9 to 23.5%-points. The effects of educational experience on both crime and enrollment also increase and become statistically significant (from -0.4 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

²⁶We present these results in Tables A1-A4 in the Appendix.

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 9.4 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 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 7, this is exactly what we see. When we do not allow accumulated experience and lagged decisions to enter, 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 26.8%-points. The marginal effect of social/emotional skills also increases in magnitude from -8.1 to -12.4%-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 9.4 to 21.0%-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 number of schools per person in the enrollment equation. In column 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. What we find is that by not including this excluded variable, the estimate of contemporaneous enrollment drops from 9.4 to 7.2%-points. The difference is not particularly big, but it 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 given the nature of our data, we are able 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 these 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 1 of Table A1 in the Appendix, 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 (reasoning and vocabulary score) have no effect on enrollment decisions. The point estimates are very small and insignificant. Given that cognitive ability is 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 shifted to the left, compared to the general population. The average 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 ability and enrollment in the baseline specification.

Consistent with the baseline estimates, the tests for cognitive ability 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 ability.

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 exceptions of the PSMI-Self-Reliance measure and in some instances PSMI-Work-Orientation, which have 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) in the face of opportunity can explain a major part of all criminal behavior. The fact that self-

reliance (and sometimes work orientation), which are viewed as positive traits, are associated with a higher probability of committing crime, suggests that positive 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. 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. As a result, 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, we estimate three other model specifications under different assumptions about enrollment. 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 the decision of whether to attend school.²⁷ Finally, we also add the indicator 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 columns 2, 3 and 4 of 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 be due to 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

²⁷ We also add the jail variable to the crime equation since the crime choice may be different when the individual is incarcerated. For instance, an incarcerated individual may be less able to commit property crimes relative to violent or drug-related crimes.

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 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 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 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 Defining Enrollment

In our baseline model we define someone as enrolled in school if they are currently 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).²⁸ The results are reported in column 5 of Table A1. Our main results are largely unchanged.

For some variables, the marginal effects estimated in the baseline shrink towards zero (female, non-intact family, and cognitive ability), and for a few other variables (age and lagged enrollment), the magnitude drops considerably. This suggests that these variables are more important in driving the decision to attend at least some school compared to none at all, than for affecting whether youths attend school on a regular basis.

For two other variables (drug use and educational experience), the effects are small and insignificant in the baseline, and in this specification increase in absolute value, and become statistically significant. Together with the baseline results, this suggests that drug use has a small effect on the decision to enroll in at least some school, but affects the intensive margin of schooling, causing people to attend less school throughout the year. With educational experience the effect is slightly different, since the new definition of enrollment not only affects the outcome (contemporaneous enrollment), but also changes how educational experience is measured. Our results indicate that just attending school for a short period of time in previous years does not generate an increase in the returns to enrollment, and that these benefits only accrue if the person attends school for most of the year.

4.3.8 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 policy effects, particularly long-run effects, may be biased. In order to examine whether this is the case, and if so to what extent, 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 6 of Table A1 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 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.

²⁸We also estimated a version of the model in which we treated enrollment in months as a continuous outcome. We estimated the system of equations as a seemingly unrelated regression (SUR) system, in which the enrollment equation is a standard linear regression, and the crime equation is a linear probability model. Although the interpretation of the results is slightly different, the results were similar qualitatively to the results for defining the cutoff to be nine months.

5 Policy Simulations

In this section we 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 drug use 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 that are 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, education policies may be an important alternative to other incapacitation policies like incarceration. If, on the other hand, other determinants of crime, skills for example, are more important, then one should consider policies that foster these skills (e.g., Heckman, 2000).

For this purpose, we present two types of simulations based on our estimated baseline model.²⁹ In the first case, we try to isolate the importance of dynamics by shocking the individuals with temporary (i.e., one year) changes, and then tracing the effects that these changes have over a period of 10 years. In particular, we simulate both the effect of not committing a crime at age 15 and the effect of not attending school at age 15. We also simulate what happens by preventing someone from consuming drugs at age 15. In the second set of simulations, we trace the dynamic effects of permanent changes to different variables that measure heterogeneity. In particular, we analyze the effects of increasing an individual’s cognitive and social/emotional skills, as well as the effect of keeping someone from consuming drugs completely.

5.1 Dynamic Effects of Temporary Changes

We begin by simulating the effects of preventing someone from 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 increases by 1.7%-points after 5 periods (when the median person has a 40% chance of being enrolled), 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 than on education. Mechanically the difference in the probability of committing a crime between the control and treatment group is one at implementation. After one year the probability of committing a crime decreases by 20%-points, when

²⁹We focus on the effects on overall crime for the median individual, since they are generally very similar to the effects on different kinds of crime, but we point out when they are different.

the median person has a 70% chance of committing a crime. 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, a person who was prevented from committing a crime at age 15 will still be around 6%-points less likely to commit a crime.

Next, we simulate the effect of having an individual enroll 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 when the policy is implemented. 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 at implementation. The effect rapidly decreases, and it reaches zero by year 3. After that it becomes slightly negative but very small as more and more human capital (i.e., years of education) gets accumulated. The effect, however, is not the same for all crimes. For the case of property crimes, enrollment has an incapacitation effect on crime and hence the effect of one more year of education at age 15 is to reduce crime by 3%-points as can be seen in Figure 7.

The third simulation we present, the effect of eliminating drug consumption at age 15, is shown in Figure 8. The effect of a one year reduction in drug consumption on enrollment is negligible. At most (after five years) it increases the probability by less than half of one percentage point. Its effect on crime, on the other hand, is quite large. At implementation, eliminating drug consumption reduces the probability of committing a crime by almost 25%-points. While the effect decreases rapidly, 10 years after implementation there is still a 2%-point lower probability of committing a crime associated with the policy.

Overall, temporary interventions lead to effects that 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. While this implies that policies based on temporary interventions will have only small effects on behavior many years after the policy (and thus may have 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 Changes

We next consider the effects that permanent changes to heterogeneity (while holding all other characteristics at their median values) may have on both the enrollment and crime probabilities. We begin by simulating the case of increasing an individual's cognitive skills from the 25th percentile to the 75th percentile in the data. While this may sound like a large change, one has to remember that this is for individuals in our selected data where this distribution is much more compressed than in the overall population. Such a change is roughly equivalent to increasing a 15 year old's IQ from around 89 to 98, or moving him from getting a score of 39 to a score of 48 on the Stroop word test, a very modest increase.³⁰ Figure 9 shows the effect of such a change on enrollment. Not surprisingly, increasing cognitive skills increases the probability of being enrolled but the effect is small. At the most it increases by 3%-points (after five years). As shown in Figure 9 increasing cognitive skills essentially has no effect on the probability of crime and, if anything, it may increase it slightly.

Figure 10 shows the effect of increasing an individual's social/emotional skills from the 25th to the 75th percentile. This change is equivalent to a 1/3 of a standard deviation change in impulse control for example. As can be seen from the figures, the effect that this change has on enrollment is negligible. A different story arises when we look at the effect of this change in social/emotional skills on criminal activity. This change reduces the probability of committing a crime by 3%-points when implemented, and the effect keeps growing over time. After 10 years the probability of committing a crime is reduced by 10%-points.

The final policy we simulate is shown in Figure 11. In this case we simulate the effects of having someone not consume drugs permanently. The effects on enrollment look very similar to the case of a temporary drug consumption change, but the effects are larger. After five years the policy increases the probability of enrollment by almost 2.5%-points, a little less than the increase due to a permanent change in cognitive skills. Its effect on crime is quite large. At implementation it reduces the probability of crime by almost 25%-points and the effects gets larger over time. After ten years the effect stabilizes at a reduction of about 45%-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

³⁰In order for one Word T-score score to be considered "higher" or "lower" than another, a 10 point or greater T 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. We find that family crime, drug use, and social/emotional skills are important drivers of criminal behavior.

In our policy simulations we illustrate the effect of individual heterogeneity on crime through policies targeting drug consumption, as the estimates of their direct effect are large. However, variables such as the probability of apprehension, future outlook, or social/emotional skills may be easier to affect. Such interventions would have similar effects as they all enter into the model in the same way, but with different magnitudes depending on the size of the policy change and the characteristics that were targeted.

We find, perhaps unsurprisingly, that policies designed to generate permanent or long-run changes, such as keeping a youth off of drugs entirely, generate the largest effects. However, policies with temporary changes to individual behavior, such as one-time reductions in drugs use, or 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 generates an effect on crime in the following year (-18%-points) that is comparable to the first year effect of permanently preventing drug use (-23%-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 (by incarceration for example) 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 comes not from 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 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.

References

- Agnew, R., T. Brezina, J. Wright, and F. Cullen. 2002. "Strain, personality traits, and delinquency: Extending general strain theory." *Criminology* 40:43–72.
- Akerlund, D., B. Golsteyn, H. Gronqvist, and L. Lindahl. 2014. "Time Preferences and Criminal Behaviour." Tech. Rep. 8168, IZA.
- Anderson, D. 2014. "In School and Out of Trouble: The Minimum Dropout Age and Juvenile Crime." *The Review of Economics and Statistics* 96(2):318–331.
- Bachman, R., R. Paternoster, and S. 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, G. S. and C. B. Mulligan. 1997. "The Endogenous Determination of Time Preference." *The Quarterly Journal of Economics* 112(3):729–758.
- Betts, J. R. and L. McFarland. 1995. "Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments." *Journal of Human Resources* 30:741–765.
- Blumstein, A., D. P. Farrington, and S. 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.
- Borghans, L., H. Meijers, and B. Ter Weel. 2008. "The role of noncognitive skills in explaining cognitive test scores." *Economic Inquiry* 46:2–12.
- Carneiro, Pedro, Karsten 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, A., T. Moffitt, P. Silva, M. Stouthamer-Loeber, R. Kreuger, and P. Schmutte. 1994. “Are some people crime prone? replication of the person-crime relationship across countries, genders, race, and methods.” *Criminology* 32:163–195.
- Cawley, John, James J. Heckman, and Edward J. Vytlačil. 2001. “Three Observations on Wages and Measured Cognitive Ability.” *Labour Economics* 8 (4):419–442.
- Chen, M.K. 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, N. 2010. “Time preference, noncognitive skills and well being across the life course: Do noncognitive skills encourage healthy behavior?” *American Economic Review: Papers and Proceedings* 100:200–204.
- Cobb-Clark, D. and M. Tan. 2011. “Noncognitive skills, occupational attainment, and relative wages.” *Labour Economics* 18:1–13.
- 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, F. and J. Heckman. 2007. “The Technology of Skill Formation.” *American Economic Review* 97(2):31–47.
- 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*, edited by Eric A. Hanushek and Frank Welch. Amsterdam: North-Holland, 697–812.
- Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. 2010. “Estimating the Technology of Cognitive and Noncognitive Skill Formation.” *Econometrica* 78 (3):883–931.
- Davis, M.L. 1988. “Time and Punishment: an Intertemporal Model of Crime.” *Journal of Political Economy* 96(2):383–390.
- Dellas, H. and P. Sakellaris. 2003. “On the Cyclicity of Schooling: Theory and Evidence.” *Oxford Economic Papers* 55:148–172.

- Dornbush, S., J Carlsmith, S. Bushwall, P. Ritter, H. Leiderman, A. Hastorf, and R. Gross. 1985. "Single parents, extended households, and the control of adolescents." *Child Development* 56(2):326–41.
- Fajnzylber, P., D. Lederman, and N. Loayza. 2002. "Inequality and violent crime." *Journal of Law and Economics* 45(1):1–40.
- Farrington, D. 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, D. P., B. Gallagher, L. Morley, R. St. Ledger, and D. West. 1986. "Unemployment, School Leaving and Crime." *British Journal of Criminology* 26(4):335–56.
- Freeman, R. 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, A., I. Garfinkel, C.E. Cooper, and R.B. Mincy. 2009. "Parental incarceration and child well-being: Implications for human beings." *Social Science Quarterly* 90(5):1186–1202.
- Gottfredson, M. and T. Hirschi. 1990. *A General Theory of Crime*. University of California Press.
- Gould, E. D., B. A. Weinberg, and D.B. Mustard. 2002. "Crime Rates and Local Labor Market Opportunities in the United States: 1977-1997." *Review of Economics and Statistics* 84(1):45–61.
- Grogger, J. 1995. "The Effect of Arrests on the Employment and Earnings of Young Men." *The Quarterly Journal of Economics* 110:51–71.
- Grogger, Jeff. 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.
- . 2000. "Policies to Foster Human Capital." *Research in Economics* 54 (1):3–56. With discussion.
- 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, P., J. Guryan, K. Roberts, K. Sixkiller, and J. Grogger. 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, T. and M. Gottfredson. 1983. "Age and the Explanation of Crime." *American Journal of Sociology* 89:552–584.
- Imai, S., H. Katayama, and K. Krishna. 2006. "Crime and Young men: The role of arrest, criminal experience and heterogeneity." Tech. Rep. 12221, NBER.
- Imai, S. and K. Krishna. 2004. "Employment, Deterrence and Crime in a Dynamic Model." *International Economic Review* 45:845–72.
- Jacob, B. and L. Lefgren. 2003. "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime." *American Economic Review* 93:1560–1577.
- Jaeger, D., T. Dohmen, A. Falk, D. Huffman, U. Sunde, and H. Bonin. 2010. "Direct evidence on risk attitudes and migration." *The Review of Economics and Statistics* 92(3):684–689.
- Kim, Dongwoo. 2014. *The Effect of Violent Crime on Educational Attainment*. Ph.D. thesis, University of Illinois at Urbana-Champaign.
- Kotlarski, Ignacy I. 1967. "On Characterizing the Gamma and Normal Distribution." *Pacific Journal of Mathematics* 20:69–76.
- Levitt, Steven. 1998. "Juvenile Crime and Punishment." *Journal of Political Economy* 106:1156–1185.
- Lochner, L. 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.
- . 2011. "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.
- Lochner, Lance J. 2004. "Education, Work, and Crime: A Human Capital Approach." *International Economic Review* 45 (3):811–843.

- 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, T.A., H. Nguyen, A.R. Piquero, and J. Fagan. 2013. "The Returns to Criminal Capital." *American Sociological Review* 78(6):925–948.
- Luallen, J. 2006. "School's Out...Forever: A Study of Juvenile Crime, At-Risk Youths and Teacher Strikes." *Journal of Urban Economics* 59:75–103.
- Merlo, A. and K. I. Wolpin. 2009. "The Transition from School to Jail: Youth Crime and High School Completion among Black Males." Penn Institute for Economic Research Working Paper 09-002, University of Pennsylvania, Philadelphia.
- Mocan, H.N. and D. Rees. 2005. "Economic Conditions, Deterrence and Juvenile Crime: Evidence from Micro Data." *American Law and Economics Review* 7(2):319–349.
- Monahan, K., L. Steinberg, E. Cauffman, and E. Mulvey. 2009. "Trajectories of Antisocial Behavior and Psychosocial Maturity from Adolescence to Young Adulthood." *Developmental Psychology* 45(6):1654–1668.
- Mulvey, E. and C. Schubert. 2012. "Some Initial Policy Implications from the Pathways to Desistance study." *Victims and Offenses* 7(4):407–427.
- 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, D., D. P. Farrington, and T. Moffitt. 1995. "Life-Course Trajectories of Different Types of Offenders." *Criminology* 33:111–140.
- Nagin, D. and K. Land. 1993. "Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model." *Criminology* 31:327–362.
- Nagin, D. and R. Paternoster. 1991. "On the relationship Of Past to Future Participation in Delinquency." *Criminology* 29:163–189.
- Nagin, D.S. and R. Paternoster. 2000. "Population Heterogeneity and State Dependence: State of the Evidence and Directions for Future Research." *Journal of Quantitative Criminology* 16:117–145.
- O'Donoghue, T. and Rabin. 1999. "Doing it Now or Later." *American Economic Review* 89(1):103–124.

- Paternoster, R. and S. Simpson. 1996. "Sanction Threats and Appeals to Morality: Testing a Rational Choice Model of Corporate Crime." *Law and Society Review* 30(3):549–83.
- Pogarsky, Greg and Alex Piquero. 2003. "Can Punishment Encourage Offending?" *Crime and Delinquency* 40(1):95–120.
- Pratt, T. and Cullen. 2000. "The empirical status of gottfredson and hirschi's general theory of crime: A meta-analysis." *Criminology* 38:931.
- Raphael, S. and R. 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):319–339.
- Schubert, C.A., E. Mulvey, L. Steinberg, E. Cauffman, S. Losoya, T. Hecker, and L. Chassin. 2004. "Operational Lessons from the Pathways to Desistance Project." *Youth Violence and Juvenile Justice* 2(3):237–255.
- Snyder, H. N. and M. Sickmund. 1999. "Juvenile offenders and victims: 1999 National Report." Tech. rep., US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.
- Tamer, Elie. 2003. "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria." *Review of Economic Studies* 70 (1):147–165.
- Tauchen, H., A.D. Witte, and H. Griesinger. 1994. "Criminal Deterrence: Revisiting the Issue with a Birth Cohort." *Review of Economics and Statistics* 76:399–412.
- Usher, D. 1997. "Education as a Deterrent to Crime." *Canadian Journal of Economics* 30(2):367–384.
- Wilson, J. Q. and R. J. Herrnstein. 1985. *Crime and Human Nature*. Simon and Schuster.
- Witte, A. 1980. "Estimating the Economic Model of Crime with Individual Data." *The Quarterly Journal of Economics* 94:57–84.
- Witte, A. and H. Tauchen. 1994. "Work and Crime: And Exploration Using Panel Data." *Public Finance* 49:1–14.
- Zeilekha, Y. and Ohad Bar-Efrat. 2009. "Crime, Terror and Corruption and Their Effect on Private Investment in Israel." *Open Economies Review* 1:1–14.

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

Variable	Drug-Related Crime		Violent Crime		Property Crime		All Crime	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Age First Crime	13.887	1.678	10.747	1.998	11.507	2.205	10.429	1.805
Age First Interview	16.026	1.141	16.029	1.141	16.030	1.140	16.029	1.141
Phoenix	0.499	0.500	0.492	0.500	0.491	0.500	0.493	0.500
Hispanic	0.342	0.475	0.339	0.474	0.339	0.474	0.340	0.474
Black	0.398	0.490	0.402	0.490	0.401	0.490	0.401	0.490
Other	0.047	0.212	0.048	0.214	0.048	0.214	0.048	0.214
Female	0.140	0.348	0.142	0.349	0.142	0.349	0.142	0.350
Siblings	4.085	2.410	4.093	2.413	4.090	2.409	4.094	2.413
Non-Intact Family	0.854	0.353	0.854	0.353	0.854	0.353	0.854	0.353
Children	0.438	0.807	0.447	0.817	0.447	0.817	0.445	0.815
Family Crime	0.195	0.396	0.193	0.395	0.193	0.394	0.195	0.396
Certainty of Punishment	5.586	2.327	5.578	2.322	5.575	2.325	5.578	2.320
Drug use	0.466	0.499	0.473	0.499	0.473	0.499	0.472	0.499
Local Unemployment Rate	5.783	1.557	5.811	1.554	5.817	1.551	5.804	1.555
Future of Outlook Inventory	2.592	0.544	2.591	0.543	2.593	0.544	2.592	0.542
Crime Rate	0.213	0.410	0.435	0.496	0.285	0.451	0.538	0.499
Enrollment Rate	0.541	0.498	0.539	0.499	0.539	0.498	0.540	0.498
Years of Education	11.102	1.968	11.122	1.965	11.117	1.966	11.116	1.967
Age	19.073	2.530	19.089	2.530	19.084	2.529	19.084	2.532
Sample Size	7210		7424		7422		7376	

Notes:

1. The descriptive statistics reported in this table correspond to data from the combined baseline and follow-up surveys.
2. The sample size varies across the four samples since they differ in the number of missing values in self-reported crime.

**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: Pathways to Desistance Descriptive Statistics:
Measures of Social/Emotional Skills**

WAI	
	% Score < 3
Impulse Control	40.3
Suppression of Aggression	50.3
Consideration of Others	17.8
PSMI	
	% Score < 2.5
Consideration of Others	5.4
Identity	4.2
Work Orientation	15.3

Notes:

1. The descriptive statistics are based on the overall crime sample.
2. The 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. Individuals are given a set of questions and asked to indicate to extent the statment is true or false on a scale of 1-5. Responses are coded such that higher numbers indicate more positive behavior. Scores below 3 indicate undesirable behavior.
3. The 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). Individuals are given a set of questions and asked to what extent they either agree or disagree with the statement on a scale of 1-4. Responses are coded such that higher numbers indicate more positive behavior. Scores below 2.5 indicate undesirable behavior.

Table 4: Distribution of Skills

	Estimate
Variance of Cognitive Skills	0.1138
Variance of Social/Emotional Skills	0.1875
Correlation Coefficient	0.2275

Notes:

We estimate a two factor model with cognitive and social/emotional measures. The table presents the estimates for the distribution of the factors, what we call cognitive and social/emotional skills, which are assumed to follow a bivariate normal distribution.

**Table 5: Estimated Parameters from Factor Analysis:
Cognitive Skills**

	WASI		Color	Stroop		Trail-Making	
	Matrix	Vocabulary		Word	Color/Word	A	B
Constant	-0.1332	0.7407	-0.4474	-1.3592	-0.9075	-	-
Age	0.0100	-0.0590	0.0763	0.1195	0.1285	-0.1345	-0.1031
Hispanic	-0.3704	-0.6158	-0.2352	-0.2763	-0.2721	0.3139	0.4108
Black	-0.6516	-0.7763	-0.4371	-0.2428	-0.4811	0.7512	0.6749
Other	-0.2577	-0.4707	-0.2360	-0.3750	-0.4350	0.2536	0.3020
Female	-0.0285	-0.0091	0.1863	0.1059	0.0674	-0.1405	-0.2419
Siblings	-0.0204	-0.0358	-0.0278	-0.0240	-0.0338	-0.0159	0.0072
Future Outlook	-0.0829	0.0286	0.0287	-0.0174	-0.0151	0.0666	0.1685
Years of Education	0.0709	0.0884	-0.0563	-0.0273	-0.0782	-0.1009	-0.0992
Cognitive Ability	1.0000	1.2244	2.0893	2.2756	1.8505	-1.3106	-1.8063
Variance	0.8235	0.7026	0.4723	0.3823	0.5466	1.0000	1.0000
Cutoff 1	-	-	-	-	-	-2.8288	-2.2188
Cutoff 2	-	-	-	-	-	-1.6384	-1.3634
Cutoff 3	-	-	-	-	-	-0.8936	-0.2749

Notes:

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: $M_{j,i}^{cog} = Z_i \beta_j^{cog} + \theta_i^{cog} \delta_j^{cog} + \xi_{j,i}^{cog}$, 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 $M_{j,i}^{cog} = 1(\psi_l < Z_i \beta_j^{cog} + \theta_i^{cog} \delta_j^{cog} + \xi_{j,i}^{cog} \leq \psi_{l+1})$.

Table 6: Estimated Parameters from Factor Analysis - Social/Emotional Skills

		WAI			PSMI		
		Impulse Control	Suppression of Aggression	Consideration of Others	Self Reliance	Identity	Work Orientation
Constant	Period 1	-1.4934	-0.7074	-2.1671	-0.6365	-0.4860	-1.2707
	Period 2	-1.3404	-0.6757	-2.2684	-0.5820	-0.4492	-1.1539
	Period 3	-1.3854	-0.6833	-2.2467	-0.5474	-0.4385	-1.0395
	Period 4	-1.3167	-0.6552	-2.2110	-0.4452	-0.3722	-0.9717
	Period 5	-1.2565	-0.5854	-2.1219	-0.3709	-0.3413	-0.8584
	Period 6	-1.2444	-0.5408	-2.0583	-0.2982	-0.2552	-0.8071
	Period 7	-1.2029	-0.4875	-2.0495	-0.2450	-0.1866	-0.7470
	Period 8	-1.1600	-0.4904	-2.0804	-0.2869	-0.2937	-0.8330
	Age in Period 1	0.0136	-0.0151	0.0142	-0.0160	-0.0297	-0.0164
	Hispanic	0.1339	-0.1075	-0.0433	-0.3039	-0.2933	-0.1851
	Black	0.4635	-0.1845	-0.0178	0.1078	0.0484	0.0015
	Other	0.2403	-0.0446	-0.0037	-0.1886	-0.1517	-0.1072
	Female	0.2133	0.1574	0.1784	0.1661	-0.0030	-0.0585
	Siblings	-0.0121	-0.0042	0.0081	-0.0106	-0.0183	-0.0025
	FOI	0.3107	0.2240	0.7484	0.1629	0.2449	0.3742
Behavioral Ability	Years of Education	0.0048	0.0328	-0.0041	0.0313	0.0311	0.0311
	Period 1	1.0000	0.9197	0.2343	1.2809	1.3246	1.2627
	Period 2	1.0324	0.9066	0.0950	1.3371	1.2880	1.2813
	Period 3	1.1184	1.0251	0.0797	1.3088	1.3425	1.2839
	Period 4	1.2610	1.1306	0.1033	1.2888	1.3373	1.2863
	Period 5	1.2910	1.1186	0.0628	1.3655	1.4156	1.3289
	Period 6	1.2720	1.1006	0.0629	1.3484	1.4266	1.4234
	Period 7	1.3725	1.1542	0.0996	1.3421	1.3833	1.4001
	Period 8	1.2565	1.1108	0.0684	1.1588	1.1932	1.1238
Variance		0.6044	0.7244	0.8050	0.5767	0.5655	0.5194

Notes:

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: $M_{ij}^{cog} = Z_i \beta_j^{cog} + \theta_i^{cog} \delta_j^{cog} + \xi_{ij}^{cog}$, where k indexes the measure (column in the table), i the individual and t time.

**Table 7: Average Marginal Effects from Probits for Crime and Education
(Overall Crime)**

VARIABLES	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting	
	(1)		(2)		(3)		(4)		(5)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Phoenix	0.0446** (0.0211)	0.0308 (0.0198)	0.0381** (0.0187)	0.0301* (0.0182)	0.0451** (0.0211)	0.0385** (0.0192)	0.0470** (0.0215)	0.0325 (0.0208)	0.0912*** (0.0184)	0.0334* (0.0199)
Hispanic	-0.0244* (0.0146)	-0.0211 (0.0154)			-0.0248* (0.0146)	-0.0232 (0.0154)	-0.0253* (0.0151)	-0.0342** (0.0158)	-0.0258* (0.0146)	-0.0218 (0.0155)
Black	0.0222 (0.0172)	-0.0353* (0.0182)			0.0217 (0.0172)	-0.0340* (0.0183)	0.0397** (0.0177)	-0.0553*** (0.0184)	0.0220 (0.0173)	-0.0349* (0.0182)
Other	0.0342 (0.0270)	-0.0246 (0.0301)			0.0340 (0.0270)	-0.0227 (0.0302)	0.0385 (0.0280)	-0.0361 (0.0303)	0.0356 (0.0270)	-0.0240 (0.0302)
Female	0.0577*** (0.0147)	-0.1020*** (0.0162)	0.0540*** (0.0144)	-0.0873*** (0.0171)	0.0574*** (0.0147)	-0.0995*** (0.0162)	0.0699*** (0.0145)	-0.1690*** (0.0156)	0.0570*** (0.0148)	-0.1020*** (0.0162)
Non-intact Family	-0.0500*** (0.0146)	0.0308* (0.0158)			-0.0504*** (0.0146)	0.0274* (0.0157)	-0.0519*** (0.0152)	0.0406** (0.0160)	-0.0509*** (0.0147)	0.0297* (0.0158)
Siblings	-0.0020 (0.0024)	0.0040 (0.0025)			-0.0018 (0.0024)	0.0039 (0.0025)	-0.0037 (0.0024)	0.0063** (0.0025)	-0.0021 (0.0024)	0.0040 (0.0025)
Age	-0.0798*** (0.0040)	-0.0253*** (0.0074)	-0.0833*** (0.0037)	-0.0347*** (0.0083)	-0.0800*** (0.0040)	-0.0345*** (0.0042)	-0.104*** (0.0020)	-0.0142 (0.0087)	-0.0819*** (0.0040)	-0.0283*** (0.0077)
Punish	0.0029 (0.0025)	-0.0221*** (0.0026)			0.0029 (0.0025)	-0.0220*** (0.0026)	0.0049* (0.0026)	-0.0279*** (0.0027)	0.0028 (0.0025)	-0.0221*** (0.0026)
Children	-0.0177** (0.0074)	0.0079 (0.0073)			-0.0174** (0.0074)	0.0067 (0.0073)	-0.0318*** (0.0076)	0.0141* (0.0074)	-0.0173** (0.0074)	0.0075 (0.0074)
Family Crime	0.0023 (0.0146)	0.1490*** (0.0152)			0.0020 (0.0146)	0.1500*** (0.0152)	-0.0019 (0.0152)	0.1760*** (0.0160)	0.0038 (0.0147)	0.150*** (0.0152)
Drug Use	-0.0022 (0.0116)	0.2240*** (0.0104)			-0.0020 (0.0116)	0.2250*** (0.0103)	-0.0114 (0.0118)	0.2680*** (0.0118)	-0.0013 (0.0116)	0.2250*** (0.0104)
Unemployment Rate	0.0208*** (0.0059)	0.0103* (0.0054)	0.0212*** (0.0059)	0.0103* (0.0059)	0.0211*** (0.0059)	0.0138*** (0.0049)	0.0226*** (0.0060)	0.0088 (0.0059)	0.0370*** (0.0047)	0.0115** (0.0055)
Future Outlook Inventory	0.0189* (0.0109)	-0.0243** (0.0115)			0.0190* (0.0109)	-0.0235** (0.0115)	0.0252** (0.0113)	-0.0313*** (0.0116)	0.0167 (0.0110)	-0.0241** (0.0115)
Years of Crime	-0.0069*** (0.0025)	0.0205*** (0.0028)	-0.0071*** (0.0023)	0.0385*** (0.0029)	-0.0069*** (0.0025)	0.0201*** (0.0028)			-0.0072*** (0.0025)	0.0204*** (0.0028)
Lagged Crime		0.1590*** (0.0124)		0.2350*** (0.0129)		0.1610*** (0.0124)				0.1600*** (0.0124)
Years of Education	0.0065 (0.0042)	-0.0043 (0.0042)	0.0114*** (0.0041)	-0.0135*** (0.0045)	0.0066 (0.0042)	-0.0025 (0.0040)			0.0057 (0.0042)	-0.0038 (0.0042)
Lagged Enrollment	0.1880*** (0.0123)		0.1910*** (0.0123)		0.1890*** (0.0123)				0.1890*** (0.0123)	
Enrollment		0.0943* (0.0484)		0.0829 (0.0535)		0.0248* (0.0142)		0.2100*** (0.0612)		0.0716 (0.0507)
Change in Schools per Young Person	322.2000*** (71.3200)		323.4000*** (71.6300)		311.6000*** (71.3100)		319.8000*** (72.2100)			
Cognitive Factor	0.0395** (0.0190)	0.0117 (0.0197)			0.0409** (0.0189)	0.0150 (0.0196)	0.0474** (0.0196)	0.0238 (0.0202)	0.0393** (0.0190)	0.0128 (0.0197)
Social/Emotional Factor	0.0026 (0.0142)	-0.0805*** (0.0148)			0.0023 (0.0142)	-0.0805*** (0.0149)	0.0124 (0.0144)	-0.1240*** (0.0149)	0.0035 (0.0142)	-0.0806*** (0.0149)
Rho	-0.1560 (0.1060)		-0.1370 (0.1020)				-0.3970*** (0.1420)		-0.1040 (0.1090)	
Observations	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190

Notes:

- 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.
- The errors in the enrollment and crime equations are allowed to be correlated in every specification, except for specification (3). Rho accounts for the correlation in errors.
- Every specification includes an exclusion restriction that enters the education equation only (Change in Schools per Young Person) except for the last specification (5).
- 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 last specification (5) does not include the exclusion restriction.

**Table 8: Average Marginal Effects from Probits for Crime and Education
(Drug-Related Crime)**

VARIABLES	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting	
	(1)		(2)		(3)		(4)		(5)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Phoenix	0.0401* (0.0212)	-0.0182 (0.0153)	0.0355* (0.0189)	-0.0174 (0.0143)	0.0408* (0.0213)	-0.0102 (0.0149)	0.0402* (0.0217)	-0.0112 (0.0159)	0.0874*** (0.0185)	-0.0165 (0.0153)
Hispanic	-0.0241* (0.0146)	-0.0223* (0.0117)			-0.0224 (0.0146)	-0.0242** (0.0117)	-0.0292* (0.0151)	-0.0287** (0.0122)	-0.0250* (0.0147)	-0.0227* (0.0117)
Black	0.0231 (0.0174)	-0.0090 (0.0140)			0.0235 (0.0174)	-0.0066 (0.0140)	0.0418** (0.0178)	-0.0257* (0.0146)	0.0231 (0.0174)	-0.0086 (0.0140)
Other	0.0337 (0.0275)	-0.0147 (0.0245)			0.0354 (0.0275)	-0.0113 (0.0245)	0.0390 (0.0286)	-0.0222 (0.0251)	0.0354 (0.0275)	-0.0142 (0.0245)
Female	0.0598*** (0.0143)	-0.1030*** (0.0143)	0.0557*** (0.0140)	-0.0985*** (0.0153)	0.0591*** (0.0143)	-0.1010*** (0.0143)	0.0665*** (0.0147)	-0.1420*** (0.0147)	0.0594*** (0.0144)	-0.1030*** (0.0143)
Non-intact Family	-0.0528*** (0.0148)	0.0329*** (0.0127)			-0.0531*** (0.0148)	0.0290** (0.0126)	-0.0552*** (0.0154)	0.0375*** (0.0131)	-0.0536*** (0.0149)	0.0321** (0.0127)
Siblings	-0.0022 (0.0024)	0.0000 (0.0019)			-0.0020 (0.0024)	-0.0001 (0.0019)	-0.0037 (0.0025)	0.0016 (0.0020)	-0.0024 (0.0024)	0.0000 (0.0019)
Age	-0.0810*** (0.0039)	0.0007 (0.0053)	-0.0840*** (0.0037)	-0.0089 (0.0057)	-0.0811*** (0.0039)	-0.0094*** (0.0032)	-0.1040*** (0.0020)	0.0097* (0.0051)	-0.0831*** (0.0039)	-0.0011 (0.0056)
Punish	0.0034 (0.0025)	-0.0103*** (0.0021)			0.0036 (0.0025)	-0.0100*** (0.0021)	0.0056** (0.0026)	-0.0148*** (0.0022)	0.0033 (0.0026)	-0.0103*** (0.0021)
Children	-0.0162** (0.0075)	0.0084 (0.0058)			-0.0161** (0.0075)	0.0072 (0.0058)	-0.0329*** (0.0077)	0.0189*** (0.0059)	-0.0158** (0.0075)	0.0082 (0.0058)
Family Crime	0.0007 (0.0148)	0.0842*** (0.0100)			0.0009 (0.0148)	0.0851*** (0.0100)	-0.0053 (0.0153)	0.1070*** (0.0105)	0.0026 (0.0148)	0.0844*** (0.0100)
Drug Use	0.0048 (0.0119)	0.2130*** (0.0095)			0.0042 (0.0119)	0.2150*** (0.0096)	-0.0070 (0.0119)	0.2560*** (0.0092)	0.0056 (0.0119)	0.2140*** (0.0096)
Unemployment Rate	0.0202*** (0.0060)	0.0018 (0.0044)	0.0207*** (0.0060)	0.0024 (0.0047)	0.0207*** (0.0060)	0.0056 (0.0040)	0.0224*** (0.0061)	0.0011 (0.0045)	0.0367*** (0.0047)	0.0025 (0.0044)
Future Outlook Inventory	0.0182 (0.0111)	-0.0082 (0.0089)			0.0191* (0.0111)	-0.0068 (0.0089)	0.0227** (0.0114)	-0.0097 (0.0094)	0.0163 (0.0111)	-0.0079 (0.0089)
Years of Crime	-0.0110*** (0.0036)	0.0218*** (0.0029)	-0.0119*** (0.0034)	0.0403*** (0.0032)	-0.0107*** (0.0036)	0.0212*** (0.0029)			-0.0113*** (0.0036)	0.0217*** (0.0029)
Lagged Crime		0.0976*** (0.0102)		0.1490*** (0.0113)		0.0984*** (0.0103)				0.0980*** (0.0102)
Years of Education	0.0067 (0.0043)	-0.0085*** (0.0033)	0.0113*** (0.0042)	-0.0115*** (0.0035)	0.0067 (0.0043)	-0.0065** (0.0032)			0.0058 (0.0043)	-0.0082** (0.0033)
Lagged Enrollment	0.1900*** (0.0125)		0.1940*** (0.0125)		0.1910*** (0.0125)				0.1910*** (0.0125)	
Enrollment		0.0785** (0.0353)		0.0634* (0.0376)		0.0001 (0.0111)		0.1550*** (0.0381)		0.0637* (0.0376)
Change in Schools per Young Person	325.5000*** (71.6600)		324.3000*** (72.0600)		310.7000*** (72.0600)		322.1000*** (72.7600)			
Cognitive Factor	0.0500*** (0.0192)	0.0138 (0.0151)			0.0495** (0.0192)	0.0182 (0.0150)	0.0642*** (0.0199)	0.0203 (0.0158)	0.0495** (0.0193)	0.0146 (0.0151)
Social/Emotional Factor	0.0018 (0.0141)	-0.0257** (0.0116)			0.0018 (0.0141)	-0.0255** (0.0116)	0.0058 (0.0146)	-0.0410*** (0.0120)	0.0030 (0.0142)	-0.0257** (0.0116)
Rho	-0.2950** (0.1300)		-0.1980* (0.1090)				-0.5530*** (0.1390)		-0.2370* (0.1360)	
Observations	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074

Notes:

- 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.
- The errors in the enrollment and crime equations are allowed to be correlated in every specification, except for specification (3). Rho accounts for the correlation in errors.
- Every specification includes an exclusion restriction that enters the education equation only (Change in Schools per Young Person) except for the last specification (5).
- 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 last specification (5) does not include the exclusion restriction.

**Table 9: Average Marginal Effects from Probits for Crime and Education
(Violent Crime)**

VARIABLES	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting	
	(1)		(2)		(3)		(4)		(5)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Phoenix	0.0390* (0.0211)	0.0263 (0.0201)	0.0321* (0.0186)	0.0360** (0.0181)	0.0395* (0.0211)	0.0346* (0.0195)	0.0435** (0.0216)	0.0156 (0.0208)	0.0868*** (0.0184)	0.0287 (0.0202)
Hispanic	-0.0220 (0.0146)	-0.0193 (0.0154)			-0.0229 (0.0146)	-0.0218 (0.0155)	-0.0260* (0.0152)	-0.0225 (0.0157)	-0.0235 (0.0146)	-0.0201 (0.0155)
Black	0.0246 (0.0172)	-0.0373** (0.0182)			0.0240 (0.0172)	-0.0358* (0.0183)	0.0404** (0.0177)	-0.0499*** (0.0184)	0.0245 (0.0173)	-0.0370** (0.0182)
Other	0.0381 (0.0271)	-0.0147 (0.0302)			0.0368 (0.0270)	-0.0124 (0.0304)	0.0405 (0.0282)	-0.0115 (0.0302)	0.0393 (0.0271)	-0.0141 (0.0303)
Female	0.0559*** (0.0147)	-0.0801*** (0.0168)	0.0519*** (0.0144)	-0.0711*** (0.0172)	0.0556*** (0.0147)	-0.0773*** (0.0169)	0.0695*** (0.0145)	-0.149*** (0.0161)	0.0550*** (0.0148)	-0.0794*** (0.0169)
Non-intact Family	-0.0477*** (0.0146)	0.0137 (0.0158)			-0.0476*** (0.0146)	0.0102 (0.0158)	-0.0511*** (0.0151)	0.0209 (0.0161)	-0.0485*** (0.0146)	0.0128 (0.0159)
Siblings	-0.0024 (0.0023)	0.00412 (0.0025)			-0.00232 (0.0023)	0.00397 (0.0025)	-0.00383 (0.0024)	0.00644** (0.0025)	-0.00259 (0.0024)	0.00408 (0.0025)
Age	-0.0806*** (0.0039)	-0.0272*** (0.0074)	-0.0839*** (0.0036)	-0.0319*** (0.0080)	-0.0805*** (0.0039)	-0.0373*** (0.0041)	-0.1040*** (0.0020)	-0.0165* (0.0085)	-0.0826*** (0.0039)	-0.0300*** (0.0077)
Punish	0.0032 (0.0025)	-0.0195*** (0.0027)			0.0032 (0.0025)	-0.0193*** (0.0027)	0.0051** (0.0026)	-0.0239*** (0.0027)	0.0031 (0.0025)	-0.0194*** (0.0027)
Children	-0.0174** (0.0073)	0.00806 (0.0075)			-0.0172** (0.0073)	0.00674 (0.0075)	-0.0319*** (0.0075)	0.0109 (0.0075)	-0.0171** (0.0074)	0.0077 (0.0075)
Family Crime	0.0045 (0.0146)	0.1300*** (0.0145)			0.0046 (0.0146)	0.1320*** (0.0145)	-0.001 (0.0152)	0.1550*** (0.0151)	0.0065 (0.0146)	0.1310*** (0.0145)
Drug Use	-0.0055 (0.0114)	0.1590*** (0.0111)			-0.0054 (0.0114)	0.1600*** (0.0111)	-0.0147 (0.0117)	0.1910*** (0.0116)	-0.0048 (0.0115)	0.1600*** (0.0111)
Unemployment Rate	0.0207*** (0.0059)	0.0046 (0.0056)	0.0210*** (0.0059)	0.0039 (0.0059)	0.0210*** (0.0059)	0.0085* (0.0051)	0.0224*** (0.0060)	0.0035 (0.0059)	0.0374*** (0.0047)	0.0057 (0.0056)
Future Outlook Inventory	0.0184* (0.0109)	-0.0249** (0.0115)			0.0177 (0.0109)	-0.0240** (0.0116)	0.0258** (0.0112)	-0.0290** (0.0117)	0.0160 (0.0109)	-0.0247** (0.0116)
Years of Crime	-0.0070*** (0.0023)	0.0210*** (0.0026)	-0.0075*** (0.0022)	0.0331*** (0.0027)	-0.0069*** (0.0023)	0.0205*** (0.0026)			-0.0073*** (0.0023)	0.0209*** (0.0026)
Lagged Crime		0.1430*** (0.0123)		0.1880*** (0.0127)		0.1450*** (0.0123)				0.1440*** (0.0123)
Years of Education	0.0067 (0.0042)	-0.0005 (0.0043)	0.0118*** (0.0041)	-0.0097** (0.0044)	0.0066 (0.0042)	0.0015 (0.0041)			0.0058 (0.0042)	0.0000 (0.0043)
Lagged Enrollment	0.1890*** (0.0123)		0.1910*** (0.0124)		0.1900*** (0.0123)				0.1900*** (0.0123)	
Enrollment		0.1090** (0.0477)		0.1010** (0.0511)		0.0335** (0.0142)		0.2040*** (0.0593)		0.0881* (0.0504)
Change in Schools per Young Person	331.5000*** (71.1300)		334.1000*** (71.4100)		320.8000*** (71.1800)		326.3000*** (72.4000)			
Cognitive Factor	0.0453** (0.0189)	0.0225 (0.0198)			0.0469** (0.0189)	0.0268 (0.0198)	0.0531*** (0.0195)	0.0407** (0.0203)	0.0452** (0.0190)	0.0237 (0.0199)
Social/Emotional Factor	0.0024 (0.0141)	-0.0727*** (0.0149)			0.0030 (0.0141)	-0.0728*** (0.0150)	0.0105 (0.0144)	-0.1100*** (0.0149)	0.0037 (0.0141)	-0.0728*** (0.0149)
Rho	-0.1680 (0.1040)		-0.1570 (0.1020)				-0.3620*** (0.1350)		-0.1200 (0.1080)	
Observations	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, except for specification (3). Rho accounts for the correlation in errors.
3. Every specification includes an exclusion restriction that enters the education equation only (Change in Schools per Young Person) except for the last specification (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 last specification (5) does not include the exclusion restriction.

**Table 10: Average Marginal Effects from Probits for Crime and Education
(Property Crime)**

VARIABLES	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting	
	(1)		(2)		(3)		(4)		(5)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Phoenix	0.0495** (0.0212)	0.0521*** (0.0172)	0.0463** (0.0189)	0.0289* (0.0156)	0.0493** (0.0212)	0.0502*** (0.0165)	0.0471** (0.0217)	0.106*** (0.0187)	0.0973*** (0.0185)	0.0536*** (0.0172)
Hispanic	-0.0233 (0.0147)	-0.0135 (0.0126)			-0.0232 (0.0147)	-0.0131 (0.0125)	-0.0264* (0.0151)	-0.0384*** (0.0133)	-0.0246* (0.0147)	-0.0139 (0.0126)
Black	0.0217 (0.0173)	-0.0103 (0.0153)			0.0217 (0.0173)	-0.0106 (0.0153)	0.0409** (0.0176)	-0.0401** (0.0162)	0.0216 (0.0173)	-0.0099 (0.0153)
Other	0.0352 (0.0272)	0.0156 (0.0248)			0.0353 (0.0272)	0.0151 (0.0247)	0.0402 (0.0283)	-0.0061 (0.0262)	0.0362 (0.0272)	0.0162 (0.0248)
Female	0.0661*** (0.0142)	-0.0256* (0.0140)	0.0624*** (0.0139)	-0.0285** (0.0143)	0.0662*** (0.0142)	-0.0265* (0.0138)	0.0693*** (0.0145)	-0.0589*** (0.0152)	0.0659*** (0.0142)	-0.0248* (0.0140)
Non-intact Family	-0.0475*** (0.0146)	0.007 (0.0131)			-0.0474*** (0.0146)	0.0078 (0.0130)	-0.0495*** (0.0152)	0.00844 (0.0141)	-0.0484*** (0.0146)	0.0064 (0.0132)
Siblings	-0.0029 (0.0024)	0.0024 (0.0021)			-0.0029 (0.0024)	0.0025 (0.0021)	-0.0040 (0.0025)	0.0030 (0.0022)	-0.0031 (0.0024)	0.0024 (0.0021)
Age	-0.0823*** (0.0038)	-0.0249*** (0.0063)	-0.0856*** (0.0036)	-0.0296*** (0.0065)	-0.0822*** (0.0038)	-0.0228*** (0.0035)	-0.104*** (0.0020)	-0.0252*** (0.0093)	-0.0846*** (0.0038)	-0.0266*** (0.0065)
Punish	0.0036 (0.0025)	-0.0157*** (0.0023)			0.0036 (0.0025)	-0.0158*** (0.0023)	0.0054** (0.0026)	-0.0220*** (0.0024)	0.0035 (0.0025)	-0.0157*** (0.0023)
Children	-0.0174** (0.0074)	0.0007 (0.0065)			-0.0176** (0.0074)	0.0010 (0.0065)	-0.0323*** (0.0076)	0.0036 (0.0070)	-0.0170** (0.0074)	0.0004 (0.0066)
Family Crime	0.0013 (0.0146)	0.0954*** (0.011)			0.0013 (0.0146)	0.0952*** (0.0112)	-0.0015 (0.0152)	0.1240*** (0.0117)	0.0032 (0.0147)	0.0955*** (0.0112)
Drug Use	-0.0075 (0.0114)	0.1440*** (0.0094)			-0.0073 (0.0114)	0.1440*** (0.0094)	-0.0138 (0.0118)	0.1830*** (0.0097)	-0.0069 (0.0115)	0.1440*** (0.0094)
Unemployment Rate	0.0214*** (0.0059)	0.0093* (0.0048)	0.0217*** (0.0059)	0.0083* (0.0050)	0.0213*** (0.0059)	0.0086** (0.0043)	0.0236*** (0.0061)	0.0119** (0.0056)	0.0378*** (0.0047)	0.0101** (0.0048)
Future Outlook Inventory	0.0176 (0.0109)	-0.0330*** (0.0098)			0.0175 (0.0109)	-0.0333*** (0.0098)	0.0240** (0.0113)	-0.0422*** (0.0105)	0.0155 (0.0109)	-0.0328*** (0.0098)
Years of Crime	-0.0034 (0.0024)	0.0184*** (0.0023)	-0.0042* (0.0022)	0.0300*** (0.0023)	-0.0034 (0.0024)	0.0184*** (0.0023)			-0.0036 (0.0024)	0.0184*** (0.0023)
Lagged Crime		0.1450*** (0.0102)		0.1960*** (0.0105)		0.1450*** (0.0102)				0.1450*** (0.0102)
Years of Education	0.0072* (0.0042)	0.0045 (0.0037)	0.0124*** (0.0041)	-0.0008 (0.0038)	0.0071* (0.0042)	0.0041 (0.0035)			0.0065 (0.0042)	0.0049 (0.0037)
Lagged Enrollment	0.1900*** (0.0123)		0.1920*** (0.0123)		0.1900*** (0.0123)				0.1900*** (0.0123)	
Enrollment		-0.0176 (0.0402)		-0.0193 (0.0414)		-0.0025 (0.0119)		-0.0032 (0.0746)		-0.0307 (0.0418)
Change in Schools per Young Person	326.6000*** (71.3900)		327.6000*** (71.6000)		328.5000*** (71.2300)		309.5000*** (74.1400)			
Cognitive Factor	0.0429** (0.0189)	0.0170 (0.0166)			0.0426** (0.0189)	0.0163 (0.0165)	0.0562*** (0.0196)	0.0395** (0.0177)	0.0426** (0.0190)	0.0176 (0.0166)
Social/Emotional Factor	0.0020 (0.0143)	-0.0660*** (0.0129)			0.0021 (0.0143)	-0.0661*** (0.0129)	0.0091 (0.0145)	-0.126*** (0.0130)	0.0031 (0.0144)	-0.0659*** (0.0129)
Rho	0.0459 (0.1170)		0.0143 (0.1070)				0.0022 (0.1950)		0.0855 (0.1220)	
Observations	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232

Notes:

- 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.
- The errors in the enrollment and crime equations are allowed to be correlated in every specification, except for specification (3). Rho accounts for the correlation in errors.
- Every specification includes an exclusion restriction that enters the education equation only (Change in Schools per Young Person) except for the last specification (5).
- 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 last specification (5) does not include the exclusion restriction.

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



Notes:

1. The figures are based on the overall crime category.
2. The horizontal axis represents the crime choice in the previous year: 1 indicates the person did engage in criminal activities last year; 0 indicates he did not.
3. The box above each figure indicates age.

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



Notes:

1. The figures are based on the overall crime category.
2. The horizontal axis represents the education choice in the previous year: 1 indicates the person enrolled in school last year; 0 indicates he did not.
3. The box above each figure indicates age.

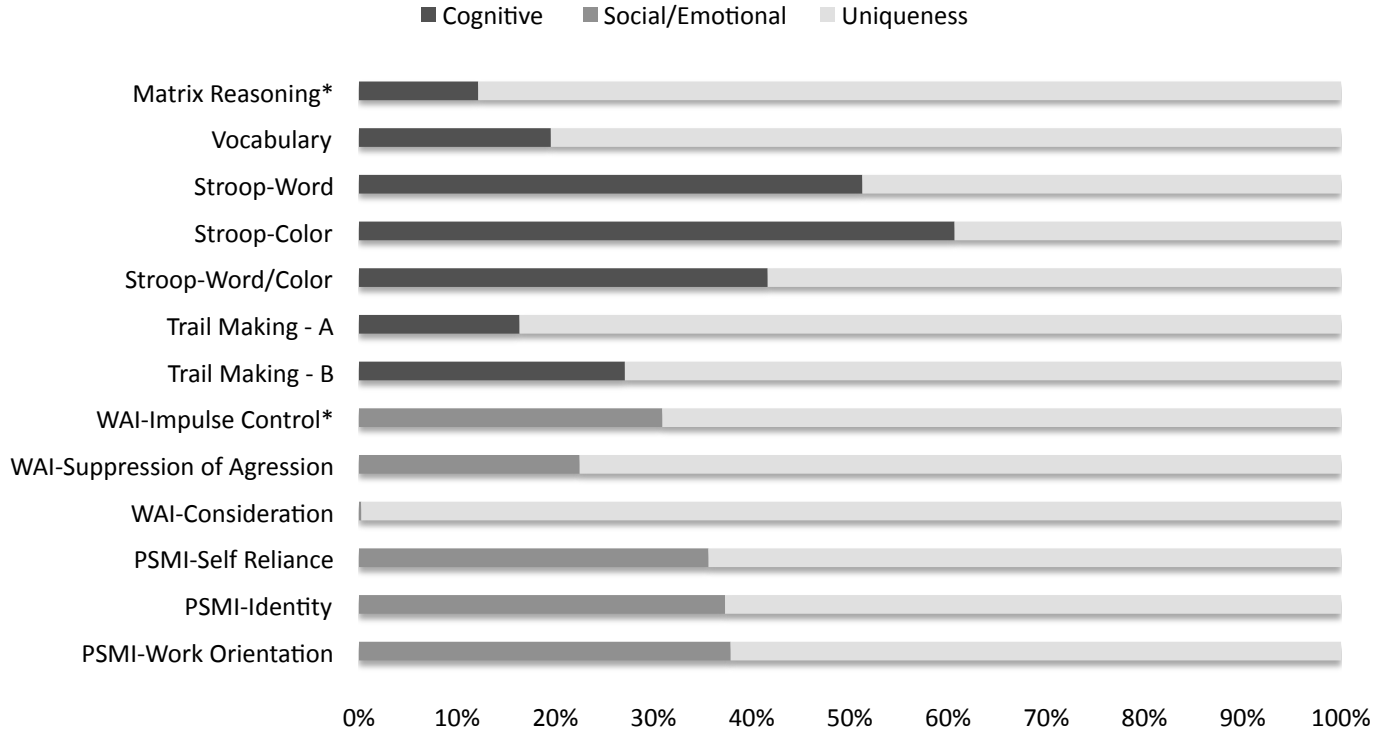
Figure 3: Probability of Crime by Enrollment Status and Age



Notes:

1. The figures are based on the overall crime category.
2. The horizontal axis represents the education choice: No indicates the person is not enrolled in school this year; Yes indicates he is.
3. The box above each figure indicates age.

Figure 4: Fraction of the Variance Explained by Skills



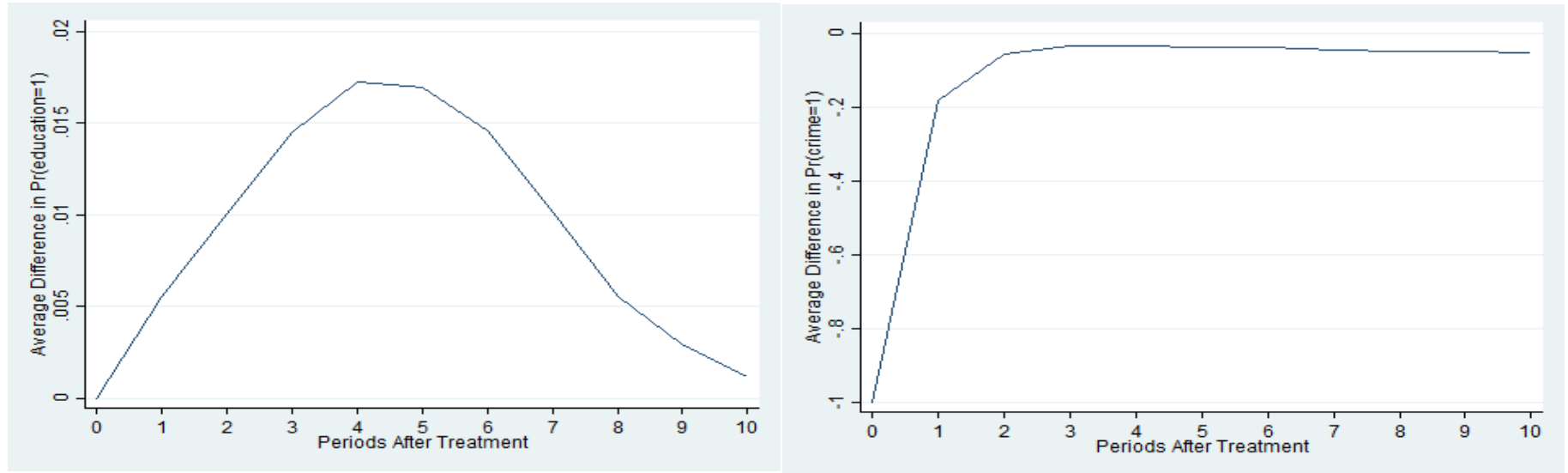
Notes:

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: $M_{j,i}^{cog} = Z_i \beta_j^{cog} + \theta_i^{cog} \delta_j^{cog} + \xi_{j,i}^{cog}$, where j indexes the measure and i the individual. For the case of Trail Making we use an ordered model of the form

$M_{j,i}^{cog} = 1(\psi_l < Z_i \beta_j^{cog} + \theta_i^{cog} \delta_j^{cog} + \xi_{j,i}^{cog} \leq \psi_{l+1})$. For the social/emotional measures we use a linear in parameters specification of the form: $M_{k,i}^{emo} = Z_{i,j} \beta_{k,j}^{emo} + \theta_i^{emo} \delta_{k,j}^{emo} + \xi_{k,i,j}^{emo}$. The figure plots the fraction of the variance explained by skills. For example, for matrix reasoning the fraction of the variance explained by cognitive skills is given by

$$\frac{\text{var}(\theta^{cog})(\delta^{cog})^2}{\text{var}(\theta^{cog})(\delta^{cog})^2 + \text{var}(\xi^{cog})}$$

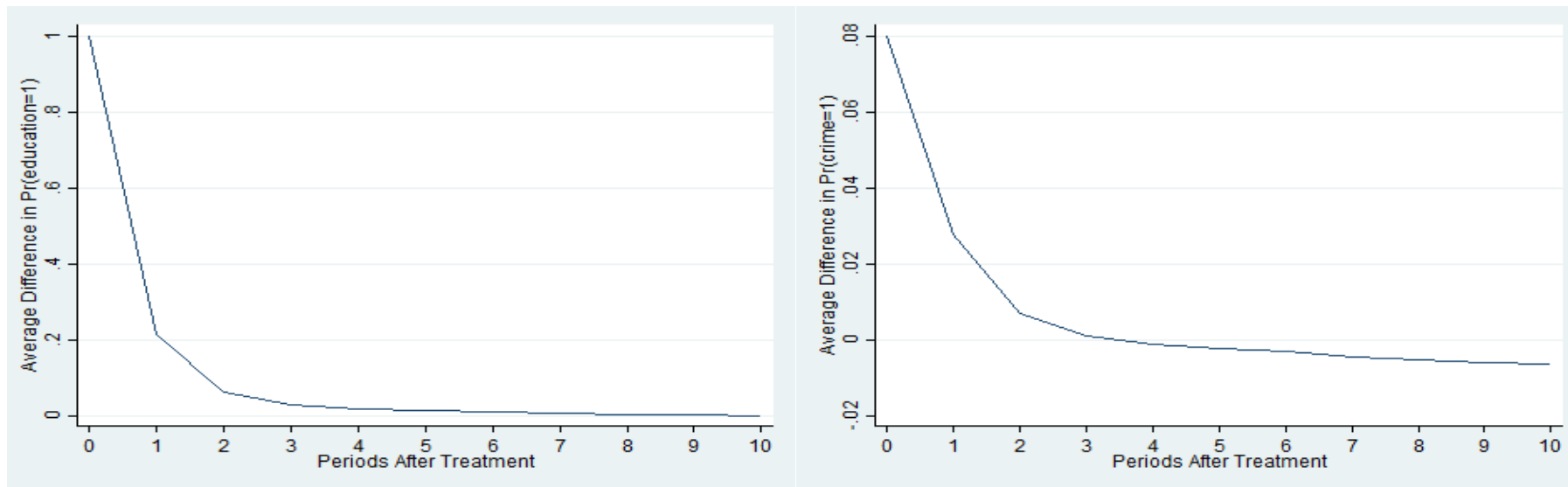
Figure 5: No Crime at Age 15 - Effect on Average Probability of Education and Crime



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 treatment of preventing someone from committing crime implies that the average difference in the probability of crime between treated and non treated is equal to -1 in the year of intervention by construction.

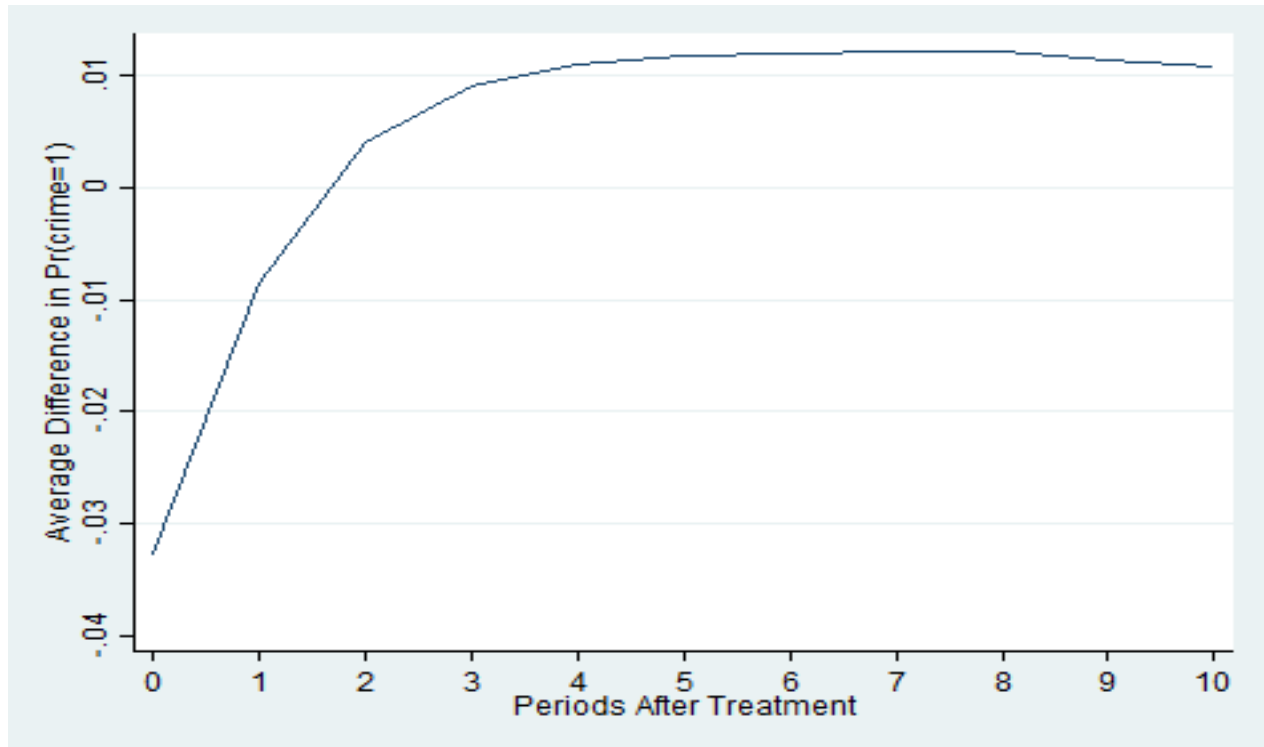
Figure 6: Enrolled at Age 15 - Effect on Average Probability of Education and Crime



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 treatment of forcing enrollment in school implies that the average difference in the probability of education between treated and non treated is equal to 1 in the year of intervention by construction.

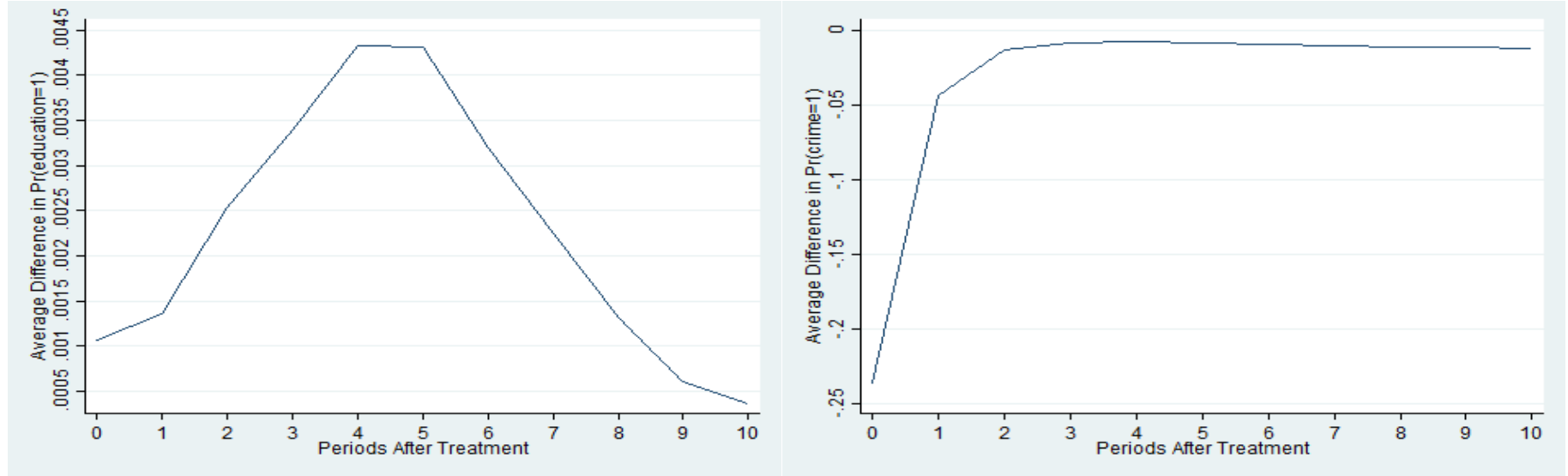
Figure 7: Enrolled at Age 15 - Effect on Average Probability of Property Crime



Notes:

1. The figure is based on the property 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.

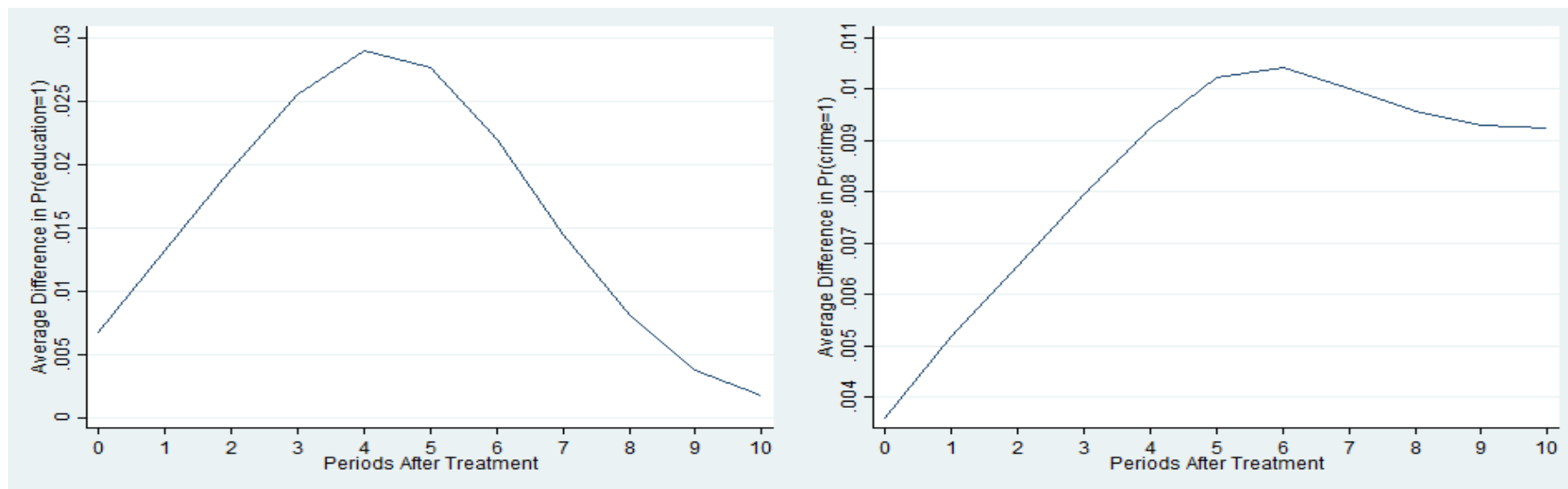
Figure 8: No Drug Use at Age 15 - Effect on Average Probability of Education and Crime



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.

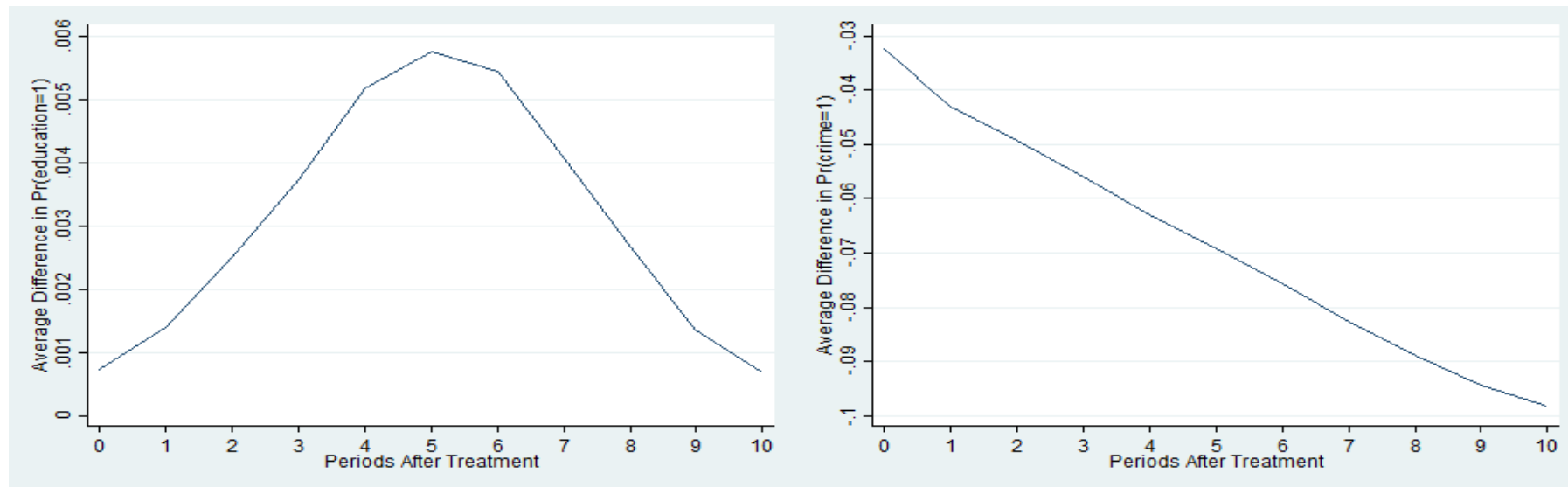
Figure 9: Move Cognitive Factor from 25th to 75th Percentile - Effect on Average Probability of Education and Crime



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.

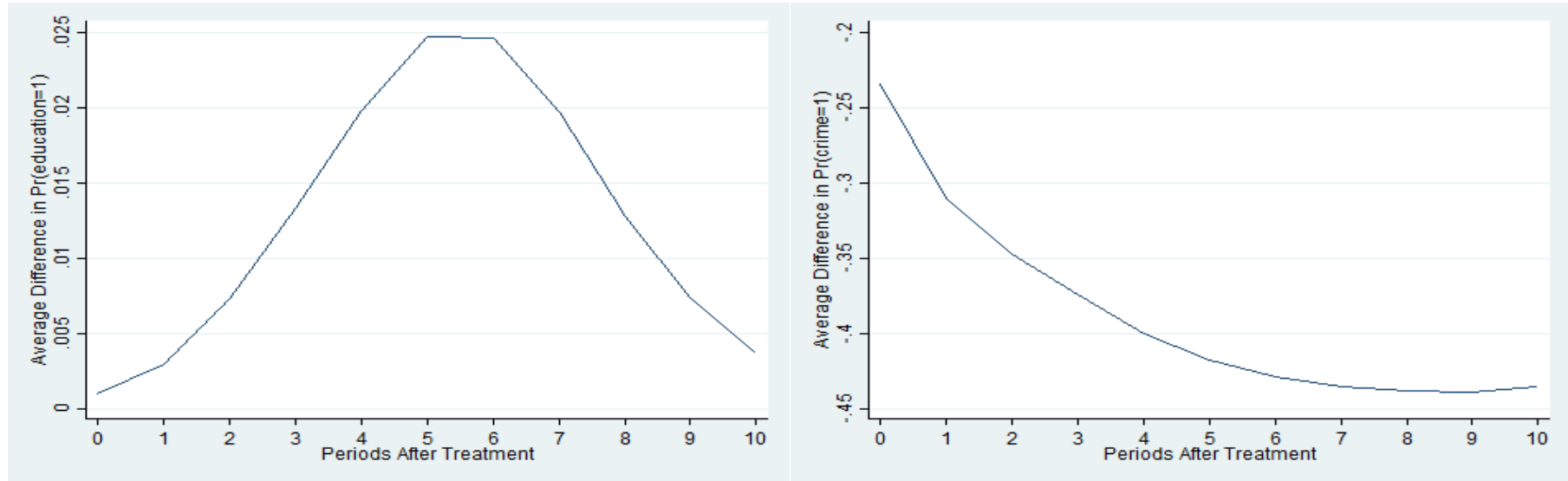
Figure 10: Move Social/Emotional Factor from 25th to 75th Percentile - Effect on Average Probability of Education and Crime



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 11: No Drug Use (Permanently) - 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 A1: Average Marginal Effects from Probits for Crime and Education (Overall Crime)

VARIABLES	Cognitive and Behavioral Ability		Choices while in Jail (1)		Choices while in Jail (2)		Choices while in Jail (3)		Enrollment based on attendance		Age varying Coefficients	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Phoenix	0.0395* (0.0216)	0.0389* (0.0202)	0.0678*** (0.0213)	0.0392** (0.0180)	0.0542*** (0.0195)	0.0290 (0.0186)	0.0526*** (0.0195)	0.0273 (0.0186)	0.0175 (0.0179)	0.0348* (0.0195)	0.0470** (0.0211)	0.0315 (0.0199)
Hispanic	-0.0129 (0.0153)	-0.0239 (0.0159)	-0.0344** (0.0145)	-0.0268* (0.0146)	-0.0055 (0.0140)	-0.0341** (0.0144)	-0.0055 (0.0139)	-0.0345** (0.0144)	-0.0210* (0.0118)	-0.0192 (0.0156)	-0.0242* (0.0146)	-0.0220 (0.0155)
Black	0.0394** (0.0180)	-0.0291 (0.0187)	-0.0014 (0.0169)	-0.0373** (0.0170)	0.0408** (0.0163)	-0.0471*** (0.0169)	0.0382** (0.0163)	-0.0480*** (0.0169)	0.0023 (0.0134)	-0.0333* (0.0183)	0.0213 (0.0172)	-0.0349* (0.0182)
Other	0.0420 (0.0272)	-0.0154 (0.0299)	0.0086 (0.0268)	-0.0202 (0.0282)	0.0324 (0.0260)	-0.0260 (0.0279)	0.0308 (0.0260)	-0.0273 (0.0279)	-0.0216 (0.0220)	-0.0215 (0.0303)	0.0348 (0.0270)	-0.0237 (0.0302)
Female	0.0529*** (0.0150)	-0.0961*** (0.0163)	0.1180*** (0.0149)	-0.0973*** (0.0157)	0.0184 (0.0150)	-0.0710*** (0.0158)	0.0182 (0.0150)	-0.0682*** (0.0159)	0.0071 (0.0116)	-0.0962*** (0.0163)	0.0592*** (0.0147)	-0.1010*** (0.0163)
Non-intact Family	-0.0486*** (0.0147)	0.0308** (0.0156)	-0.0577*** (0.0145)	0.0227 (0.0149)	-0.0301** (0.0142)	0.0155 (0.0148)	-0.0299** (0.0142)	0.0139 (0.0148)	-0.0038 (0.0118)	0.0268* (0.0157)	-0.0515*** (0.0146)	0.0294* (0.0158)
Siblings	-0.0017 (0.0024)	0.0035 (0.0025)	-0.0025 (0.0023)	0.0039* (0.0023)	-0.0013 (0.0022)	0.0034 (0.0022)	-0.0012 (0.0022)	0.0036 (0.0022)	-0.0044** (0.0019)	0.0043* (0.0025)	-0.0020 (0.0024)	0.0039 (0.0025)
Age	-0.0788*** (0.0041)	-0.0261*** (0.0072)	-0.0708*** (0.0041)	-0.0322*** (0.0040)	-0.0736*** (0.0036)	-0.0324*** (0.0068)	-0.0732*** (0.0036)	-0.0324*** (0.0068)	-0.0463*** (0.0031)	-0.0343*** (0.0051)	-0.0872*** (0.0052)	-0.0247*** (0.0077)
Punish	0.0022 (0.0026)	-0.0184*** (0.0026)	0.0066*** (0.0025)	-0.0189*** (0.0024)	-0.0009 (0.0024)	-0.0166*** (0.0024)	-0.0010 (0.0024)	-0.0169*** (0.0024)	0.0054*** (0.0020)	-0.0223*** (0.0027)	0.0027 (0.0025)	-0.0220*** (0.0026)
Children	-0.0173** (0.0074)	0.0027 (0.0073)	-0.0160** (0.0077)	0.0079 (0.0069)	-0.0151** (0.0066)	0.0116* (0.0069)	-0.0133** (0.0066)	0.0119* (0.0069)	-0.0249*** (0.0077)	0.0102 (0.0075)	-0.0173** (0.0074)	0.0074 (0.0074)
Family Crime	0.0015 (0.0146)	0.1460*** (0.0151)	-0.0335** (0.0137)	0.1440*** (0.0137)	0.0088 (0.0136)	0.1310*** (0.0137)	0.0073 (0.0135)	0.1330*** (0.0137)	0.0099 (0.0112)	0.1460*** (0.0154)	0.0024 (0.0147)	0.1490*** (0.0152)
Drug Use	-0.0011 (0.0118)	0.2040*** (0.0105)	0.0428*** (0.0111)	0.2260*** (0.0097)	-0.0405*** (0.0108)	0.2330*** (0.0096)	-0.0412*** (0.0108)	0.2320*** (0.0096)	-0.0456*** (0.0091)	0.2300*** (0.0105)	-0.0012 (0.0116)	0.2240*** (0.0104)
Unemployment Rate	0.0212*** (0.0059)	0.0097* (0.0053)	0.0200*** (0.0062)	0.0122** (0.0047)	0.0225*** (0.0054)	0.0107** (0.0052)	0.0227*** (0.0054)	0.0101* (0.0052)	0.0176*** (0.0054)	0.0120** (0.0052)	0.0214*** (0.0059)	0.0106* (0.0054)
Future Outlook Inventory	0.0225* (0.0122)	0.0160 (0.0127)	0.0192* (0.0107)	-0.0291*** (0.0106)	0.0156 (0.0102)	-0.0275*** (0.0105)	0.0158 (0.0102)	-0.0269** (0.0105)	0.0177** (0.0088)	-0.0259** (0.0116)	0.0197* (0.0109)	-0.0240** (0.0115)
Years of Crime	-0.0066** (0.0026)	0.0171*** (0.0028)	-0.0073*** (0.0025)	0.0208*** (0.0026)	-0.0045* (0.0024)	0.0200*** (0.0026)	-0.0051* (0.0027)	0.0217*** (0.0030)	-0.0051** (0.0021)	0.0209*** (0.0029)		
Lagged Crime		0.1420*** (0.0124)		0.1590*** (0.0118)		0.1500*** (0.0117)		0.1630*** (0.0136)		0.1580*** (0.0126)		
Years of Education	0.0060 (0.0042)	-0.0050 (0.0042)	0.0232*** (0.0042)	-0.0038 (0.0037)	-0.0075* (0.0038)	-0.0008 (0.0038)	-0.0024 (0.0043)	-0.0025 (0.0044)	0.0162*** (0.0035)	0.0033 (0.0040)		
Lagged Enrollment	0.1850*** (0.0124)		0.1730*** (0.0129)		0.1980*** (0.0114)		0.2230*** (0.0131)		0.0836*** (0.0099)			
Enrollment		0.0958** (0.0470)		0.0561*** (0.0157)		0.0632 (0.0465)		0.0684 (0.0461)		0.0974* (0.0563)		
Change in Schools per Young Person	310.7000*** (71.2600)		166.7000** (71.6300)		318.3000*** (67.1200)		319.4000*** (66.9600)		120.8000* (63.4000)		318.3000*** (71.6000)	
Cognitive Factor			0.0352* (0.0185)	0.0037 (0.0181)	0.0075 (0.0177)	0.0132 (0.0180)	0.0052 (0.0176)	0.0138 (0.0180)	0.0040 (0.0153)	0.0152 (0.0199)	0.0386** (0.0189)	0.0136 (0.0198)
Social/Emotional Factor			0.0152 (0.0137)	-0.0765*** (0.0137)	-0.0122 (0.0133)	-0.0714*** (0.0136)	-0.0131 (0.0132)	-0.0714*** (0.0136)	0.0040 (0.0112)	-0.0828*** (0.0150)	0.0025 (0.0142)	-0.0808*** (0.0149)
Jail					0.1000*** (0.0119)	0.1180*** (0.0131)	0.3560*** (0.0780)	0.1600* (0.0915)				
Years of Crime * Jail							0.0027 (0.0047)	-0.0066 (0.0050)				
Years of Education * Jail							-0.0196*** (0.0065)	0.0052 (0.0073)				
Lagged Enrollment * Jail							-0.0895*** (0.0231)					
Lagged Crime * Jail							-0.0521** (0.0256)					
Enrollment * Jail							-0.0311 (0.0258)					
Years of Crime * Age1											-0.0062 (0.0038)	0.0217*** (0.0044)
Years of Crime * Age2											-0.0072** (0.0029)	0.0197*** (0.0034)
Years of Education * Age1											0.0019 (0.0048)	-0.0029 (0.0050)
Years of Education * Age2											0.0092** (0.0044)	-0.0036 (0.0044)
Lagged Enrollment * Age1											0.2290*** (0.0200)	
Lagged Enrollment * Age2											0.1700*** (0.0162)	
Enrollment * Age1												0.0697 (0.0499)

Table A1 (Continued): Average Marginal Effects from Probits for Crime and Education (Overall Crime)

VARIABLES	Cognitive and Behavioral Ability		Choices while in Jail (1)		Choices while in Jail (2)		Choices while in Jail (3)		Enrollment based on attendance		Age varying Coefficients	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Enrollment * Age2											0.0533	(0.0573)
Lagged Crime * Age1											0.1570***	(0.0188)
Lagged Crime * Age2											0.1610***	(0.0167)
WASI Reasoning Score	-0.0017 (0.0064)	-0.0048 (0.0067)										
WASI Vocabulary Score	0.0004 (0.0070)	0.0011 (0.0072)										
Stroop: Color	-0.0030 (0.0078)	0.0143* (0.0082)										
Stroop: Word	0.0087 (0.0071)	-0.0116 (0.0075)										
Stroop: Color - Word	0.0040 (0.0069)	-0.0103 (0.0072)										
Trail-Making: Part A	-0.0034 (0.0067)	-0.0018 (0.0068)										
Trail-Making: Part B	-0.0164** (0.0069)	-0.0071 (0.0071)										
WAI - Impulse Response	-0.0076 (0.0074)	-0.0301*** (0.0075)										
WAI - Suppression of Aggression	0.0102 (0.0072)	-0.0445*** (0.0073)										
WAI - Consideration of Others	0.0015 (0.0062)	-0.0275*** (0.0063)										
PSMI - Self Reliance	-0.0143 (0.0104)	0.0227** (0.0106)										
PSMI - Identity	0.0355*** (0.0103)	-0.0175* (0.0106)										
PSMI - Work Orientation	-0.0225** (0.0090)	-0.0115 (0.0095)										
Rho	-0.1570 (0.1050)		-0.0704** (0.0327)		-0.0899 (0.0999)		-0.0831 (0.0992)		-0.2280* (0.1240)		-0.0858 (0.1110)	
Observations	5,190	5,190	6,189	6,189	6,189	6,189	6,189	6,189	5,097	5,097	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 replace our factor estimates of cognitive and social/emotional ability with the measures used to infer them. 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.

Table A2: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime)

VARIABLES	Cognitive and Behavioral Ability		Choices while in Jail (1)		Choices while in Jail (2)		Choices while in Jail (3)		Enrollment based on attendance		Age varying Coefficients	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Phoenix	0.0329 (0.0218)	-0.0137 (0.0158)	0.0625*** (0.0215)	-0.00303 (0.0141)	0.0492** (0.0197)	-0.0097 (0.0145)	0.0482** (0.0196)	-0.0117 (0.0145)	0.0136 (0.0182)	-0.0103 (0.0150)	0.0420** (0.0212)	-0.0160 (0.0153)
Hispanic	-0.0105 (0.0154)	-0.0271** (0.0121)	-0.0319** (0.0145)	-0.0148 (0.0110)	-0.0053 (0.0140)	-0.0189* (0.0110)	-0.00519 (0.0139)	-0.0192* (0.0109)	-0.0194 (0.0118)	-0.0236** (0.0118)	-0.0238 (0.0146)	-0.0213* (0.0117)
Black	0.0435** (0.0181)	-0.0119 (0.0144)	-0.0056 (0.0171)	-0.0002 (0.0131)	0.0441*** (0.0165)	-0.0080 (0.0132)	0.0419** (0.0165)	-0.0097 (0.0132)	0.0016 (0.0136)	-0.0098 (0.0141)	0.0223 (0.0173)	-0.0093 (0.0140)
Other	0.0418 (0.0277)	-0.0103 (0.0247)	0.0123 (0.0273)	-0.0073 (0.0233)	0.0328 (0.0266)	-0.0089 (0.0231)	0.0314 (0.0265)	-0.0088 (0.0231)	-0.0219 (0.0224)	-0.0121 (0.0246)	0.0359 (0.0274)	-0.0140 (0.0245)
Female	0.0531*** (0.0146)	-0.0993*** (0.0144)	0.1140*** (0.0145)	-0.0919*** (0.0140)	0.0230 (0.0146)	-0.0780*** (0.0141)	0.0232 (0.0146)	-0.0764*** (0.0142)	0.0131 (0.0114)	-0.1000*** (0.0144)	0.0615*** (0.0143)	-0.1020*** (0.0143)
Non-intact Family	-0.0513*** (0.0149)	0.0332*** (0.0127)	-0.0614*** (0.0147)	0.0209* (0.0119)	-0.0329** (0.0144)	0.0184 (0.0118)	-0.0328** (0.0144)	0.0184 (0.0119)	-0.0066 (0.0120)	0.0299** (0.0127)	-0.0539*** (0.0148)	0.0329*** (0.0127)
Siblings	-0.0019 (0.0024)	0.0001 (0.0019)	-0.0023 (0.0023)	-0.0012 (0.0018)	-0.0019 (0.0022)	-0.0012 (0.0017)	-0.0017 (0.0022)	-0.0011 (0.0017)	-0.0044** (0.0020)	-0.0005 (0.0020)	-0.0022 (0.0024)	0.0001 (0.0019)
Age	-0.0796*** (0.0040)	0.0011 (0.0052)	-0.0698*** (0.0040)	-0.0090*** (0.0031)	-0.0755*** (0.0035)	-0.0033 (0.0050)	-0.0750*** (0.0035)	-0.0038 (0.0050)	-0.0481*** (0.0031)	-0.0121*** (0.0037)	-0.0879*** (0.0052)	0.0004 (0.0057)
Punish	0.0028 (0.0026)	-0.0079*** (0.0021)	0.0067*** (0.0025)	-0.0081*** (0.0020)	0.0000 (0.0024)	-0.0072*** (0.0020)	0.0000 (0.0024)	-0.0073*** (0.0020)	0.0061*** (0.0021)	-0.0099*** (0.002)	0.0031 (0.0025)	-0.0104*** (0.0021)
Children	-0.0158** (0.0075)	0.0074 (0.0058)	-0.0126 (0.0078)	0.0099* (0.0054)	-0.0150** (0.0067)	0.0124** (0.0054)	-0.0133** (0.0067)	0.0125** (0.0055)	-0.0244*** (0.0079)	0.0117* (0.0059)	-0.0162** (0.0075)	0.0085 (0.0058)
Family Crime	0.0000 (0.0148)	0.0816*** (0.0100)	-0.0294** (0.0139)	0.0778*** (0.0092)	0.0053 (0.0137)	0.0689*** (0.0092)	0.0036 (0.0137)	0.0689*** (0.0092)	0.0086 (0.0114)	0.0833*** (0.0101)	0.0015 (0.0148)	0.0833*** (0.0100)
Drug Use	0.0059 (0.0120)	0.2030*** (0.0095)	0.0553*** (0.0114)	0.229*** (0.0085)	-0.0384*** (0.0111)	0.2360*** (0.0085)	-0.0397*** (0.0111)	0.2360*** (0.0085)	-0.0450*** (0.0094)	0.2170*** (0.0102)	0.0060 (0.0118)	0.2120*** (0.0095)
Unemployment Rate	0.0206*** (0.0060)	0.0011 (0.0043)	0.0202*** (0.0062)	0.0064* (0.0039)	0.0219*** (0.0055)	0.0039 (0.0041)	0.0220*** (0.0055)	0.0039 (0.0041)	0.0177*** (0.0054)	0.0064 (0.0042)	0.0206*** (0.0060)	0.0026 (0.0044)
Future Outlook Inventory	0.0238* (0.0124)	0.0068 (0.0098)	0.0194* (0.0108)	-0.0089 (0.0083)	0.0147 (0.0103)	-0.0074 (0.0083)	0.0149 (0.0103)	-0.0081 (0.0083)	0.0173* (0.0090)	-0.0068 (0.0090)	0.0195* (0.0111)	-0.0085 (0.0089)
Years of Crime	-0.0113*** (0.0036)	0.0201*** (0.0029)	-0.0210*** (0.0036)	0.0231*** (0.0028)	-0.0018 (0.0033)	0.0211*** (0.0028)	-0.0026 (0.0039)	0.0229*** (0.0033)	-0.0033 (0.0032)	0.0212*** (0.0030)		
Lagged Crime		0.0920*** (0.0102)		0.0881*** (0.0096)		0.0859*** (0.0095)		0.0985*** (0.0118)		0.0988*** (0.0104)		
Years of Education	0.0060 (0.0043)	-0.0090*** (0.0033)	0.0222*** (0.0043)	-0.0048 (0.0030)	-0.0071* (0.0039)	-0.0033 (0.0030)	-0.0022 (0.0043)	-0.0038 (0.0036)	0.0172*** (0.0036)	-0.0022 (0.0032)		
Lagged Enrollment	0.1860*** (0.0125)		0.1780*** (0.0130)		0.1990*** (0.0115)		0.2230*** (0.0131)		0.0854*** (0.0101)			
Enrollment		0.0883** (0.0344)		-0.0020 (0.0125)		0.0436 (0.0345)		0.0396 (0.0347)		-0.0046 (0.0481)		
Change in Schools per Young Person	311.8000*** (71.4900)		143.8000** (72.3800)		323.5000*** (67.1800)		325.4000*** (67.0700)		139.8000** (65.4700)		331.1000*** (71.5900)	
Cognitive Factor			0.0425** (0.0187)	0.0081 (0.0141)	0.0128 (0.0179)	0.0148 (0.0140)	0.0105 (0.0179)	0.0141 (0.0141)	0.0085 (0.0157)	0.0212 (0.0152)	0.0483** (0.0192)	0.0123 (0.0151)
Social/Emotional Factor			0.0115 (0.0137)	-0.0232** (0.0108)	-0.0111 (0.0132)	-0.0214** (0.0107)	-0.0118 (0.0132)	-0.0218** (0.0107)	0.00776 (0.0112)	-0.0272** (0.0118)	0.0015 (0.0141)	-0.0263** (0.0116)
Jail					0.1050*** (0.0122)	0.0625*** (0.00956)	0.3660*** (0.0761)	0.0668 (0.0651)				
Years of Crime * Jail							0.0040 (0.0065)	-0.0062 (0.0056)				
Years of Education * Jail							-0.0191*** (0.0065)	0.00219 (0.0053)				
Lagged Enrollment * Jail							-0.0884*** (0.0233)					
Lagged Crime * Jail								-0.0342* (0.0200)				
Enrollment * Jail								-0.0006 (0.0189)				
Years of Crime * Age1											-0.0163*** (0.0055)	0.0272*** (0.0051)
Years of Crime * Age2											-0.00763* (0.0044)	0.0190*** (0.0035)
Years of Education * Age1											0.0037 (0.0046)	-0.0090** (0.0035)
Years of Education * Age2											0.0089** (0.0044)	-0.0089*** (0.0034)
Lagged Enrollment * Age1											0.2270*** (0.0203)	
Lagged Enrollment * Age2											0.1710*** (0.0163)	
Enrollment * Age1												0.0815** (0.0360)

Table A2 (Continued): Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime)

VARIABLES	Cognitive and Behavioral Ability		Choices while in Jail (1)		Choices while in Jail (2)		Choices while in Jail (3)		Enrollment based on attendance		Age varying Coefficients	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Enrollment * Age2											0.0918**	(0.0449)
Lagged Crime * Age1											0.0721***	(0.0149)
Lagged Crime * Age2											0.1240***	(0.0149)
WASI Reasoning Score	-0.0006	0.0008										
	(0.0064)	(0.0051)										
WASI Vocabulary Score	0.0020	-0.0016										
	(0.0070)	(0.0055)										
Stroop: Color	-0.0037	0.0159**										
	(0.0079)	(0.0063)										
Stroop: Word	0.0101	-0.0038										
	(0.0072)	(0.0061)										
Stroop: Color - Word	0.0053	-0.0071										
	(0.0070)	(0.0054)										
Trail-Making: Part A	-0.0042	0.0043										
	(0.0067)	(0.0052)										
Trail-Making: Part B	-0.0180**	0.0039										
	(0.0069)	(0.0054)										
WAI - Impulse Response	-0.0075	-0.0081										
	(0.0075)	(0.0058)										
WAI - Suppression of Aggression	0.0105	-0.0251***										
	(0.0073)	(0.0058)										
WAI - Consideration of Others	0.0000	-0.0116**										
	(0.0063)	(0.0048)										
PSMI - Self Reliance	-0.0135	0.0197**										
	(0.0105)	(0.0080)										
PSMI - Identity	0.0363***	-0.0294***										
	(0.0104)	(0.0078)										
PSMI - Work Orientation	-0.0255***	0.0097										
	(0.0091)	(0.0071)										
Rho	-0.3200**		0.0289		-0.1730		-0.1590		-0.0528		-0.3170**	
	(0.1290)		(0.0414)		(0.1200)		(0.1210)		(0.1650)		(0.1410)	
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

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 replace our factor estimates of cognitive and social/emotional ability with the measures used to infer them. 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.

Table A3: Average Marginal Effects from Probits for Crime and Education (Violent Crime)

VARIABLES	Cognitive and Behavioral Ability		Choices while in Jail (1)		Choices while in Jail (2)		Choices while in Jail (3)		Enrollment based on attendance		Age varying Coefficients	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Phoenix	0.0353 (0.0216)	0.0433** (0.0206)	0.0620*** (0.0213)	0.0321* (0.0184)	0.0508*** (0.0195)	0.0258 (0.0191)	0.0496** (0.0195)	0.0252 (0.0191)	0.0111 (0.0179)	0.0308 (0.0198)	0.0414** (0.0211)	0.0270 (0.0202)
Hispanic	-0.0105 (0.0153)	-0.0260 (0.0159)	-0.0318** (0.0144)	-0.0220 (0.0146)	-0.0050 (0.0139)	-0.0265* (0.0145)	-0.0049 (0.0139)	-0.0270* (0.0145)	-0.0177 (0.0117)	-0.0169 (0.0156)	-0.0222 (0.0146)	-0.0204 (0.0155)
Black	0.0432** (0.0180)	-0.0323* (0.0187)	0.0010 (0.0169)	-0.0396** (0.0171)	0.0417** (0.0163)	-0.0461*** (0.0171)	0.0392** (0.0163)	-0.0464*** (0.0171)	0.0064 (0.0134)	-0.0334* (0.0183)	0.0236 (0.0172)	-0.0368** (0.0183)
Other	0.0461* (0.0273)	-0.0065 (0.0302)	0.0110 (0.0268)	-0.0053 (0.0283)	0.0341 (0.0261)	-0.0071 (0.0281)	0.0321 (0.0260)	-0.0072 (0.0281)	-0.0193 (0.0221)	-0.0102 (0.0303)	0.0382 (0.0270)	-0.0127 (0.0303)
Female	0.0513*** (0.0150)	-0.0748*** (0.0169)	0.1160*** (0.0149)	-0.0794*** (0.0164)	0.0170 (0.0150)	-0.0597*** (0.0165)	0.0177 (0.0150)	-0.0584*** (0.0166)	0.0010 (0.0117)	-0.0707*** (0.0170)	0.0570*** (0.0147)	-0.0782*** (0.0170)
Non-intact Family	-0.0461*** (0.0146)	0.0153 (0.0157)	-0.0549*** (0.0145)	0.0104 (0.0150)	-0.0282** (0.0142)	0.0054 (0.0150)	-0.0284** (0.0141)	0.0044 (0.0150)	-0.0034 (0.0117)	0.0097 (0.0158)	-0.0489*** (0.0146)	0.0117 (0.0160)
Siblings	-0.0021 (0.0024)	0.0039 (0.0025)	-0.0028 (0.0023)	0.0030 (0.0023)	-0.0015 (0.0022)	0.0026 (0.0023)	-0.0014 (0.0022)	0.0026 (0.0023)	-0.0050** (0.0019)	0.0042 (0.0025)	-0.0024 (0.0023)	0.0039 (0.0025)
Age	-0.0794*** (0.0040)	-0.0286*** (0.0073)	-0.0715*** (0.0040)	-0.0340*** (0.0039)	-0.0739*** (0.0035)	-0.0359*** (0.0068)	-0.0733*** (0.0035)	-0.0361*** (0.0068)	-0.0454*** (0.0030)	-0.0352*** (0.0052)	-0.0876*** (0.0051)	-0.0284*** (0.0077)
Punish	0.0025 (0.0025)	-0.0156*** (0.0027)	0.0066*** (0.0025)	-0.0169*** (0.0025)	-0.0005 (0.0024)	-0.0151*** (0.0025)	-0.0006 (0.0023)	-0.0152*** (0.0025)	0.0058*** (0.0020)	-0.0196*** (0.0027)	0.0029 (0.0025)	-0.0194*** (0.0027)
Children	-0.0169** (0.0074)	0.0029 (0.0074)	-0.0153** (0.0076)	0.0082 (0.0071)	-0.0153** (0.0066)	0.0107 (0.0071)	-0.0136** (0.0065)	0.0104 (0.0071)	-0.0241*** (0.0077)	0.0071 (0.0077)	-0.0170** (0.0073)	0.0075 (0.0075)
Family Crime	0.0035 (0.0146)	0.1260*** (0.0144)	-0.0319** (0.0136)	0.1370*** (0.0131)	0.0110 (0.0136)	0.1260*** (0.0132)	0.0096 (0.0135)	0.1280*** (0.0132)	0.0116 (0.0111)	0.1280*** (0.0148)	0.0049 (0.0146)	0.1310*** (0.0145)
Drug Use	-0.0043 (0.0116)	0.1380*** (0.0112)	0.0396*** (0.0110)	0.1610*** (0.0103)	-0.0429*** (0.0107)	0.1650*** (0.0104)	-0.0440*** (0.0107)	0.1650*** (0.0104)	-0.0426*** (0.0090)	0.1660*** (0.0113)	-0.0045 (0.0114)	0.1600*** (0.0111)
Unemployment Rate	0.0213*** (0.0059)	0.0052 (0.0055)	0.0196*** (0.0062)	0.0065 (0.0049)	0.0227*** (0.0054)	0.0059 (0.0053)	0.0228*** (0.0054)	0.0054 (0.0054)	0.0171*** (0.0053)	0.0059 (0.0055)	0.0214*** (0.0059)	0.0050 (0.0056)
Future Outlook Inventory	0.0223* (0.0121)	0.0118 (0.0127)	0.0186* (0.0106)	-0.0308*** (0.0107)	0.0144 (0.0101)	-0.0293*** (0.0107)	0.0146 (0.0101)	-0.0287*** (0.0107)	0.0180** (0.00878)	-0.0259** (0.0117)	0.0189* (0.0109)	-0.0239** (0.0116)
Years of Crime	-0.0065*** (0.0023)	0.0170*** (0.0027)	-0.0074*** (0.0023)	0.0215*** (0.0025)	-0.0047** (0.0022)	0.0204*** (0.0025)	-0.0046* (0.0025)	0.0204*** (0.0029)	-0.0067*** (0.0019)	0.0221*** (0.0027)		
Lagged Crime		0.1250*** (0.0122)		0.1480*** (0.0115)		0.1450*** (0.0115)		0.1600*** (0.0137)		0.1420*** (0.0124)		
Years of Education	0.0062 (0.0042)	-0.0003 (0.0043)	0.0233*** (0.0043)	-0.0006 (0.0038)	-0.0072* (0.0038)	0.0021 (0.0039)	-0.0022 (0.0042)	-0.0006 (0.0045)	0.0147*** (0.0035)	0.0055 (0.0041)		
Lagged Enrollment	0.1850*** (0.0124)		0.1730*** (0.0129)		0.1990*** (0.0113)		0.2240*** (0.0130)		0.0849*** (0.0101)			
Enrollment		0.1060** (0.0470)		0.0726*** (0.0157)		0.0630 (0.0468)		0.0549 (0.0467)		0.1260** (0.0617)		
Change in Schools per Young Person	319.3000*** (71.1100)		180.9000** (71.5100)		325.9000*** (67.0200)		327.0000*** (66.8600)		126.0000** (63.5400)		326.9000*** (71.4300)	
Cognitive Factor			0.0400** (0.0184)	0.0122 (0.0183)	0.0131 (0.0176)	0.0207 (0.0183)	0.0105 (0.0176)	0.0219 (0.0183)	0.0111 (0.0152)	0.0261 (0.0200)	0.0446** (0.0189)	0.0248 (0.0199)
Social/Emotional Factor			0.0159 (0.0136)	-0.0702*** (0.0138)	-0.0120 (0.0132)	-0.0667*** (0.0138)	-0.0127 (0.0131)	-0.0663*** (0.0138)	0.0023 (0.0111)	-0.0735*** (0.0151)	0.0027 (0.0141)	-0.0732*** (0.0150)
Jail					0.1000*** (0.0119)	0.0897*** (0.0130)	0.374*** (0.0770)	0.0042 (0.0892)				
Years of Crime * Jail								0.0000 (0.0045)		-0.0011 (0.0049)		
Years of Education * Jail								-0.0195*** (0.0064)		0.0097 (0.0071)		
Lagged Enrollment * Jail								-0.0882*** (0.0230)				
Lagged Crime * Jail								-0.0499** (0.0246)				
Enrollment * Jail								0.0196 (0.0256)				
Years of Crime * Age1											-0.0070** (0.0034)	0.0195*** (0.0041)
Years of Crime * Age2											-0.0070** (0.0028)	0.0217*** (0.0033)
Years of Education * Age1											0.0023 (0.0046)	0.0025 (0.0048)
Years of Education * Age2											0.0090** (0.0044)	0.0002 (0.0045)
Lagged Enrollment * Age1											0.2280*** (0.0201)	
Lagged Enrollment * Age2											0.1710*** (0.0162)	
Enrollment * Age1												0.0836* (0.0495)

Table A3 (Continued): Average Marginal Effects from Probits for Crime and Education (Violent Crime)

VARIABLES	Cognitive and Behavioral Ability		Choices while in Jail (1)		Choices while in Jail (2)		Choices while in Jail (3)		Enrollment based on attendance		Age varying Coefficients	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Enrollment * Age2											0.0495	(0.0578)
Lagged Crime * Age1											0.1410***	(0.0176)
Lagged Crime * Age2											0.1470***	(0.0175)
WASI Reasoning Score	0.0004 (0.0064)	-0.0050 (0.0067)										
WASI Vocabulary Score	-0.0016 (0.0070)	-0.0087 (0.0073)										
Stroop: Color	-0.0023 (0.0077)	0.0081 (0.0081)										
Stroop: Word	0.0099 (0.0070)	-0.0033 (0.0075)										
Stroop: Color - Word	0.0039 (0.0069)	-0.0019 (0.0072)										
Trail-Making: Part A	-0.0034 (0.0067)	-0.0012 (0.0069)										
Trail-Making: Part B	-0.0165** (0.0068)	-0.0065 (0.0071)										
WAI - Impulse Response	-0.0085 (0.0073)	-0.0248*** (0.0075)										
WAI - Suppression of Aggression	0.0106 (0.0072)	-0.0586*** (0.0074)										
WAI - Consideration of Others	0.0013 (0.0062)	-0.0271*** (0.0063)										
PSMI - Self Reliance	-0.0142 (0.0103)	0.0126 (0.0106)										
PSMI - Identity	0.0351*** (0.0103)	-0.0080 (0.0106)										
PSMI - Work Orientation	-0.0215** (0.0090)	0.0039 (0.0093)										
Rho	-0.1590 (0.1040)		-0.0851*** (0.0321)		-0.0540 (0.0979)		-0.0513 (0.0982)		-0.2690* (0.1370)		-0.0801 (0.1090)	
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

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 replace our factor estimates of cognitive and social/emotional ability with the measures used to infer them. 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.

Table A4: Average Marginal Effects from Probits for Crime and Education (Property Crime)

VARIABLES	Cognitive and Behavioral Ability		Choices while in Jail (1)		Choices while in Jail (2)		Choices while in Jail (3)		Enrollment based on attendance		Age varying Coefficients	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Phoenix	0.0459** (0.0218)	0.0462*** (0.0175)	0.0730*** (0.0215)	0.0683*** (0.0156)	0.0586*** (0.0196)	0.0674*** (0.0162)	0.0575*** (0.0196)	0.0675*** (0.0162)	0.0178 (0.0180)	0.0523*** (0.0168)	0.0524** (0.0212)	0.0543*** (0.0172)
Hispanic	-0.0127 (0.0154)	-0.0107 (0.0129)	-0.0335** (0.0146)	-0.0248** (0.0118)	-0.0053 (0.0140)	-0.0269** (0.0118)	-0.0056 (0.0140)	-0.0270** (0.0118)	-0.0210* (0.0119)	-0.0147 (0.0127)	-0.0231 (0.0147)	-0.0133 (0.0126)
Black	0.0392** (0.0179)	-0.0008 (0.0156)	-0.0036 (0.0170)	-0.0208 (0.0143)	0.0413** (0.0163)	-0.0236 (0.0143)	0.0389** (0.0163)	-0.0232 (0.0143)	0.0013 (0.0135)	-0.0087 (0.0154)	0.0209 (0.0172)	-0.0093 (0.0153)
Other	0.0424 (0.0274)	0.0210 (0.0245)	0.0104 (0.0270)	-0.0091 (0.0236)	0.0337 (0.0263)	-0.0081 (0.0236)	0.0310 (0.0262)	-0.0087 (0.0237)	-0.0237 (0.0223)	0.0124 (0.0249)	0.0358 (0.0271)	0.0166 (0.0248)
Female	0.0601*** (0.0145)	-0.0208 (0.0140)	0.126*** (0.0143)	-0.0239* (0.0135)	0.0242* (0.0145)	-0.0135 (0.0137)	0.0250* (0.0145)	-0.0121 (0.0138)	0.0113 (0.0112)	-0.0267* (0.0139)	0.0674*** (0.0142)	-0.0232* (0.0140)
Non-intact Family	-0.0460*** (0.0146)	0.0049 (0.0130)	-0.0553*** (0.0145)	0.0031 (0.0123)	-0.0277* (0.0142)	0.0004 (0.0123)	-0.0281** (0.0141)	-0.0005 (0.0124)	-0.0029 (0.0118)	0.0082 (0.0131)	-0.0486*** (0.0146)	0.0067 (0.0131)
Siblings	-0.0026 (0.0024)	0.0024 (0.0021)	-0.0035 (0.0023)	0.0027 (0.0019)	-0.0021 (0.0022)	0.0025 (0.0019)	-0.0019 (0.0022)	0.0026 (0.0019)	-0.0045** (0.0019)	0.0025 (0.0021)	-0.0030 (0.0023)	0.0022 (0.0021)
Age	-0.0809*** (0.0039)	-0.0254*** (0.0061)	-0.0730*** (0.0039)	-0.0244*** (0.0034)	-0.0749*** (0.0035)	-0.0266*** (0.0055)	-0.0742*** (0.0035)	-0.0268*** (0.0055)	-0.0467*** (0.0030)	-0.0260*** (0.0043)	-0.0902*** (0.0051)	-0.0236*** (0.0065)
Punish	0.0027 (0.0025)	-0.0129*** (0.0023)	0.0069*** (0.0025)	-0.0134*** (0.0022)	0.0000 (0.0024)	-0.0126*** (0.0022)	-0.0002 (0.0024)	-0.0129*** (0.0022)	0.0061*** (0.0021)	-0.0154*** (0.0024)	0.0033 (0.0025)	-0.0154*** (0.0023)
Children	-0.0167** (0.0074)	-0.0023 (0.0065)	-0.0163** (0.0076)	0.0018 (0.0062)	-0.0155** (0.0066)	0.0031 (0.0062)	-0.0137** (0.0066)	0.0021 (0.0063)	-0.0234*** (0.0077)	-0.0009 (0.0066)	-0.0168** (0.0073)	0.0003 (0.0065)
Family Crime	0.0010 (0.0146)	0.0904*** (0.0111)	-0.0347** (0.0137)	0.0908*** (0.0103)	0.0085 (0.0136)	0.0856*** (0.0104)	0.0072 (0.0135)	0.0871*** (0.0103)	0.0098 (0.0112)	0.0966*** (0.0114)	0.0020 (0.0146)	0.0950*** (0.0112)
Drug Use	-0.00480 (0.0117)	0.1260*** (0.0094)	0.0389*** (0.0110)	0.1540*** (0.0086)	-0.0441*** (0.0107)	0.1560*** (0.0088)	-0.0452*** (0.0107)	0.1550*** (0.0088)	-0.0443*** (0.0091)	0.1390*** (0.0104)	-0.0066 (0.0114)	0.1430*** (0.0094)
Unemployment Rate	0.0218*** (0.0059)	0.0079* (0.0047)	0.0204*** (0.0062)	0.0118*** (0.0042)	0.0232*** (0.0054)	0.0121*** (0.0045)	0.0234*** (0.0054)	0.0121*** (0.0045)	0.0169*** (0.0054)	0.0103** (0.0046)	0.0220*** (0.0059)	0.0102** (0.0048)
Future Outlook Inventory	0.0219* (0.0122)	0.00271 (0.0109)	0.0180* (0.0106)	-0.0374*** (0.00913)	0.0133 (0.0101)	-0.0365*** (0.00913)	0.0136 (0.0101)	-0.0370*** (0.00912)	0.0166* (0.00885)	-0.0335*** (0.00992)	0.0187* (0.0109)	-0.0334*** (0.0098)
Years of Crime	-0.0034 (0.0024)	0.0149*** (0.0023)	-0.0046* (0.0024)	0.0180*** (0.0022)	-0.0024 (0.0022)	0.0180*** (0.0022)	-0.0019 (0.0025)	0.0194*** (0.0025)	-0.0037* (0.0020)	0.0185*** (0.0023)		
Lagged Crime		0.1320*** (0.0102)		0.1440*** (0.00958)		0.1410*** (0.00956)		0.1580*** (0.0117)		0.1460*** (0.0103)		
Years of Education	0.0066 (0.0042)	0.0031 (0.0037)	0.0236*** (0.0043)	0.0041 (0.0032)	-0.0073* (0.0038)	0.0058* (0.0033)	-0.0022 (0.0042)	0.0052 (0.0039)	0.0152*** (0.0035)	0.0042 (0.0035)		
Lagged Enrollment	0.1860*** (0.0123)		0.1750*** (0.0129)		0.2000*** (0.0113)		0.2230*** (0.0129)		0.0896*** (0.0098)			
Enrollment		-0.0111 (0.0396)		-0.0058 (0.0134)		-0.0175 (0.0377)		-0.0286 (0.0374)		-0.0754 (0.0494)		
Change in Schools per Young Person	314.5000*** (71.2800)		175.2000** (71.6000)		323.4000*** (66.7600)		323.6000*** (66.6200)		130.5000** (63.5600)		327.6000*** (71.5800)	
Cognitive Factor			0.0368** (0.0185)	0.0061 (0.0155)	0.0113 (0.0176)	0.0100 (0.0155)	0.0086 (0.0176)	0.0110 (0.0155)	0.0090 (0.0153)	0.0130 (0.0168)	0.0417** (0.0189)	0.0188 (0.0166)
Social/Emotional Factor			0.0134 (0.0139)	-0.0630*** (0.0119)	-0.0116 (0.0134)	-0.0617*** (0.0119)	-0.0123 (0.0134)	-0.0613*** (0.0119)	0.00379 (0.0113)	-0.0643*** (0.0130)	0.00180 (0.0143)	-0.0664*** (0.0128)
Jail					0.1000*** (0.0119)	0.0440*** (0.0106)	0.3870*** (0.0774)	0.0578 (0.0736)				
Years of Crime * Jail							-0.0017 (0.0043)	-0.0045 (0.0041)				
Years of Education * Jail							-0.0200*** (0.0064)	0.0018 (0.0058)				
Lagged Enrollment * Jail							-0.0862*** (0.0229)					
Lagged Crime * Jail								-0.0499** (0.0205)				
Enrollment * Jail								0.0201 (0.0215)				
Years of Crime * Age1											-0.0046 (0.0034)	0.0241*** (0.0033)
Years of Crime * Age2											-0.0026 (0.0028)	0.0141*** (0.0028)
Years of Education * Age1											0.0034 (0.0045)	0.0037 (0.0040)
Years of Education * Age2											0.0095** (0.0043)	0.0056 (0.0038)
Lagged Enrollment * Age1											0.2280*** (0.0199)	
Lagged Enrollment * Age2											0.1720*** (0.0162)	
Enrollment * Age1												-0.0322 (0.0405)

Table A4 (Continued): Average Marginal Effects from Probits for Crime and Education (Property Crime)

VARIABLES	Cognitive and Behavioral Ability		Choices while in Jail (1)		Choices while in Jail (2)		Choices while in Jail (3)		Enrollment based on attendance		Age varying Coefficients	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime	Education	Crime
Enrollment * Age2											-0.0456 (0.0459)	
Lagged Crime * Age1											0.1310*** (0.0140)	
Lagged Crime * Age2											0.1580*** (0.0159)	
WASI Reasoning Score	0.0006 (0.0064)	0.0011 (0.0056)										
WASI Vocabulary Score	-0.0027 (0.0070)	0.0131** (0.0060)										
Stroop: Color	-0.0005 (0.0077)	0.0046 (0.0068)										
Stroop: Word	0.0082 (0.0070)	-0.0024 (0.0063)										
Stroop: Color - Word	0.0033 (0.0069)	-0.0110* (0.0060)										
Trail-Making: Part A	-0.0035 (0.0066)	0.0002 (0.0058)										
Trail-Making: Part B	-0.0163** (0.0068)	-0.0069 (0.0059)										
WAI - Impulse Response	-0.0073 (0.0074)	-0.0302*** (0.0063)										
WAI - Suppression of Aggression	0.0127* (0.0071)	-0.0286*** (0.0064)										
WAI - Consideration of Others	0.0011 (0.0062)	-0.0220*** (0.0053)										
PSMI - Self Reliance	-0.0131 (0.0103)	0.0122 (0.0087)										
PSMI - Identity	0.0349*** (0.0102)	-0.0124 (0.0085)										
PSMI - Work Orientation	-0.0244*** (0.0090)	-0.0152** (0.0075)										
Rho	0.0408 (0.119)		0.0035 (0.0361)		0.0245 (0.1060)		0.0363 (0.1060)		0.1740 (0.1470)		0.1110 (0.1210)	
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

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 replace our factor estimates of cognitive and social/emotional ability with the measures used to infer them. 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.