9-5-2018 11:00 AM

Essays on Crime, Education, and Employment

Maria Antonella Mancino, *The University of Western Ontario*

Supervisor: Dr. Salvador Navarro, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Economics

Abstract

My thesis consists of three chapters that are motivated by policy-relevance and contribute to the study of crime choices among young individuals.

Chapter 2 studies the determinants of youth crime using a dynamic discrete choice model of crime and education. We allow past education and criminal activities to affect current crime and educational decisions. We take advantage of a rich panel dataset on serious juvenile offenders, the Pathways to Desistance. Using a series of psychometric tests, we estimate a model of cognitive and social/emotional skills which feed into the crime and education model. This allows us to separately identify the roles of state dependence, returns to experience, and heterogeneity in driving crime and enrollment decisions among youth. We find small effects of experience and stronger evidence of state dependence and heterogeneity for crime and schooling. We provide evidence that, as a consequence, policies that affect individual heterogeneity (e.g., social/emotional skills), and those that temporarily keep youth away from crime, can have important and lasting effects even if criminal experience has already accumulated.

Chapter 3 documents empirical facts about the criminal and legal labour sector for disadvantaged young individuals, and investigates the factors driving the transitions between sectors. I focus on the role of heterogeneity, earnings, human capital, and criminal capital in determining transitions across the criminal and legal labour sectors. The data I employ comes from the Pathways to Desistance Study. I find that disadvantaged young individuals face two low-quality employment alternatives. On the one hand, jobs in the legal labour sector are characterized by short average duration and low wages. Consistent with their low quality, these jobs present small returns to legal experience. Activity in the criminal sector presents similar features as legal jobs and it offers an earnings premium relative to the legal labour sector, which partially compensates for the inherent risk of the activity. I provide evidence that earnings in the criminal and legal labour sectors play a significant role on the transitions across sectors. This implies that choices in the criminal and legal labour sectors are strongly related and, as a result, they should not be studied in isolation of each other.

Motivated by the findings in Chapter 3, Chapter 4 analyzes the legal employment and crime choices for disadvantaged youth. The labour market for this population group is usually studied ignoring the presence of the criminal sector, and yet a large share of them participate in crime. To study these outcomes jointly and explore how they relate, I build and estimate a two-sector search model allowing for a rich set of interactions between the two sectors. I estimate the model using monthly data from the Pathways to Desistance. Search frictions in the legal labour sector are found to be significant, with these individuals being offered low-quality legal jobs that are characterized by low earnings and large destruction rates. The criminal sector provides an attractive alternative to the legal labour sector, offering an earnings premium. Nevertheless, crime brings a higher probability of incarceration and fewer opportunities in the legal labour sector. I find that there are sizable interactions across sectors, and that policies in one sector can have important effects on the other sector. I provide evidence that policies targeting the legal labour sector (e.g., wage subsidy) can reduce crime and boost legal employment among disadvantaged youth. Furthermore, a policy that reduces the arrival rate of crime opportunities (e.g., via increasing the number of police), compared to extending the average sentence length, has the advantage of reducing crime without generating large increases in the incarcerated population.

Keywords: Crime, Education, Employment, Youth.

Co-Authorship Statement

This thesis contains material co-authored with David Rivers and Salvador Navarro. All the authors are equally responsible for the work which appears in Chapter 2 of this thesis.

Acknowlegements

I am greatly indebted to my supervisor Salvador Navarro and my dissertation committee members, Audra Bowlus, Lance Lochner, and David Rivers, for their support and guidance in developing my research. I am grateful to Salvador Navarro and David Rivers for making my first co-authorship experience a truly great one. I further want to thank them for the opportunity to work as their research assistant, which greatly helped me develop my skills as an economist.

I would like to thank conference participants at the 2017 UM-MSU-UWO Labor Day Conference, the 2016 Canadian Economic Association Annual Meetings, and the 2016 North American Summer Meeting of the Econometric Society, for valuable comments and recommendations. I extend my gratitude to Nirav Mehta, Joseph Mullins, Tim Conley, and other seminar participants at the University of Western Ontario, for being a constant source of feedback. I would also like to extend my gratitude to my fellow graduate students, Diego Salazar, Miguel Cardoso, Qian Liu, Tom Handler, Galyna Grynkiv, Sergii Pypko, Sha Wang, Wenya Wang, Martin Luccioni, and Ali Kamranzadeh.

The completion of the dissertation would not have been possible without the generous financial support of the Department of Economics at the University of Western Ontario. I am grateful for all the support from the economics staff, Sandra Augustine, Karin Feulgen, Alexandra Houston, Jennifer Hope, Leslie Kostal, Debra Merrifield, Maureen O'Connell, and Sharon Phillips.

Finally, I am deeply grateful to my mother, Analia, and my late father, Ricardo, for their unconditional love and support of all my aspirations. My Ph.D endeavor would not be possible without the love, patience, and support from my beloved husband, Juan Ignacio. Lastly, I would not have made it this far without the affection and constant encouragement from my dear brothers, Juan Andres and Luciano, my grandparents, and parents-in-law.

To my family

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

My thesis consists of three chapters that are motivated by policy-relevance and contribute to the study of crime choices among young individuals. The first chapter, written in co-authorship with Salvador Navarro and David Rivers, studies the determinants of youth crime and education. The second chapter provides a descriptive picture of the labour market for disadvantaged youth, encompassing both criminal and legal activities. The last chapter explores the interactions between the criminal and legal labour sectors.

All three chapters take advantage of a rich panel dataset on serious juvenile offenders, the Pathways to Desistance (PDS). The PDS was designed specifically to study questions related to the evolution of criminal behaviour, taking special care to also measure educational and employment outcomes. As a result, the dataset contains a rich panel of information about decisions to participate in crime, have a legal employment, and enroll in school. Each study participant was followed for a period of seven years after entering the survey which results in a comprehensive picture of life changes in a wide array of areas over the course of this time. These features make the PDS data well-suited for understanding the dynamics in crime, legal employment, and education.

Chapter 2 studies the determinants of youth crime in the context of a joint dynamic discrete choice model of crime and education, by allowing previous decisions to affect current choices. Different from the vast majority of papers studying the relationship between crime and education, this chapter focuses on serious offenders rather than studying the population at large. For policy makers interested in reducing overall crime rates, data on these serious offenders, who contribute significantly to aggregate crime rates, is necessary. Using a series of psychometric tests designed to measure unobserved heterogeneity, we estimate a model of cognitive and social/emotional skills which feeds into the crime and education model. Furthermore, we include a large set of targeted control variables beyond what is typically available. The extremely rich set of control variables allows us to separate the effects of experience 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). Lastly, incorporating these additional measures of observed and unobserved heterogeneity also represent additional potential instruments for policy makers.

We find small effects of experience and stronger evidence of state dependence and heterogeneity for crime and schooling. 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 social/emotional skills, have some of the largest effects. As a result, policies that affect individual heterogeneity (e.g., social/emotional skills), and those that temporarily keep youth away from crime, can have important and lasting effects even if criminal experience has already accumulated.

Chapter 3 documents empirical facts about the criminal and legal labour sector for disadvantaged young individuals and investigates the factors driving the transitions between sectors. I focus on the role of heterogeneity, earnings, human capital, and criminal capital in determining transitions across the criminal and legal labour sectors. Given the large fraction of disadvantaged young individuals engaged in crime, a thorough analysis of the labour sector should encompass both criminal and legal labour sector activities.

Using data from the PDS, I find that young individuals face two low-quality employment alternatives. First, jobs in the legal labour sector are characterized by short duration and low wages, on average. These jobs present small returns to legal experience. Altogether, the findings suggest that legal jobs are of low quality. Activity in the criminal sector presents similar features as legal jobs and it offers an earnings premium relative to the legal labour sector, which partially compensates for the inherent risk of the activity. I provide evidence that earnings in the criminal and legal labour sectors play a significant role on the transitions across sectors. This implies that choices in the criminal and legal labour sectors are strongly related and, as a result, they should not be studied in isolation of each other.

Motivated by the strong correlation between activity in the criminal and legal labour sectors, Chapter 4 analyzes the legal employment and crime choices for disadvantaged youth jointly. To study these outcomes and explore how they relate, I build and estimate a two-sector search model allowing for a rich set of interactions between the two sectors. The search framework is well-suited to this type of study because it brings together some key features of the criminal and legal labour sectors documented in Chapter 3, such as the long periods of non-employment and/or criminal activity, and ties them to frictions and earnings differences in the criminal and legal labour sectors.

I find that search frictions in the legal labour sector are significant, with young individuals being offered low-quality legal jobs that are characterized by low earnings and large destruction rates. In this context where legal jobs are not easily available, the criminal sector provides an attractive alternative to the legal labour sector, offering a sizable earnings premium. Nevertheless, crime brings a higher probability of incarceration and fewer opportunities in the legal labour sector. I find that there are important interactions across sectors. For example, increasing the arrival rate of legal jobs yields a sizable reduction in crime. As a result, policies in one sector can have important effects on the other sector. I provide evidence that policies targeting the legal labour sector (e.g., wage subsidy) can reduce crime and boost legal employment among disadvantaged youth. Furthermore, a policy that reduces the arrival rate of crime opportunities (e.g., via increasing the number of police), compared to extending the average sentence length, has the advantage of reducing crime without generating large increases in the incarcerated population.

My thesis aims at understanding the main drivers of the crime choice, as well as which policies are effective in reducing crime among youth. Overall, the results suggest that policies that increase education or that improve the quality and the access to the legal labour sector can have important effects on crime. In the next three chapters, I provide more detailed information about each of these findings.

Chapter 2

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

2.1 Introduction

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

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

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

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

frontations (Jacob and Lefgren, 2003) or increased drug-related offenses by bringing together buyers and sellers.[3] Schooling can affect social networks, and these networks could influence criminal behaviour, for example via gang participation (Lochner, 2010).

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

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

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

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

An advantage of studying only serious offenders through the PDS data is that, to the extent that there is unobserved heterogeneity that leads some individuals to become serious offenders, we are more likely to be observing individuals who are on a criminal trajectory (Nagin and Land, 1993; Nagin, Farrington, and Moffitt, 1995). For policy makers interested in reducing

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

[4]We describe the dataset in more detail in Section 2.2.

overall crime rates, particularly violent crime rates, data on these serious offenders, who contribute significantly to aggregate crime rates, is necessary. While selecting on serious offenders has its advantages, one limitation is that we cannot necessarily generalize our findings to the population at large. The data are also less useful for studying the transition to becoming a serious offender, as we only observe those individuals that have already done so.

Our extremely rich set of control variables allows us to separate the effects of experience (captured by the accumulation of education and crime) from contemporaneous effects of education on crime, and from the effects of individual heterogeneity. Furthermore, we are able to separately account for the effects of state dependence in these decisions (captured by lagged decisions). Being able to separate 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 PDS data includes a much larger set of targeted control variables than is typically available. In addition to standard socio-economic variables and information about individuals' families, the dataset also contains a number of additional individual-level variables that are particularly useful for our analysis. In each year the data contain a measure of each individual's perception about their probability of being caught if they commit a crime.[5] It also has information about drug usage, involvement in crime by family members, and a measure of how each individual discounts future events, among others.

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

Incorporating these additional measures of observed and unobserved heterogeneity not only aids in separately identifying the various channels driving observed crime and education decisions. They also represent additional potential instruments for policy makers. To the extent

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

that behavioural problems or drug use affect criminal activity, this provides additional opportunities to affect criminal behaviour 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 social/emotional skills, have some of the largest effects. We also find evidence of important dynamics. State dependence leads to the strongest effects, but there is evidence of small returns to experience.

The rest of the chapter is organized as follows. In Section 2.2 we describe our data from the Pathways to Desistance. Section 2.3 contains our joint dynamic discrete choice model of crime and education. Section 2.4 presents the empirical results from our model, as well as a number of robustness checks. In Section 2.5, we provide some simulations of our model to illustrate how the enrollment and criminal behaviour evolve over time and discuss some policy implications. Section 2.6 concludes.

2.2 Data

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

The initial (baseline) survey occurred when individuals first entered the sample. For those in the juvenile system, the baseline interview was completed within 75 days after their adjudication, and for those in the adult system within 90 days after their decertification hearing (in Philadelphia) or arraignment (in Phoenix). There were six semi-annual follow-up interviews,

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

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

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

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

followed by four annual follow-up interviews. They were typically conducted in the participant's home, or in a residential facility if the individual was in a jail or juvenile detention center. In total, the survey covers each individual for eight years. Individuals were paid \$50 to participate in the baseline survey, with compensation increasing for the follow-ups to minimize attrition (Monahan et al., 2009). The retention rate, measured as the share of participants completing a particular interview wave, was above 90% for the first six waves and no less than 83% for the following annual interviews.

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

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

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

In addition to the detailed information about observable characteristics of each individual, the PDS 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)

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

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

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

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

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

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

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

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

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

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

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

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

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

We construct four panel datasets, one for each of the three crime measures described above and one with all crime together. Each panel includes all individuals for whom all the relevant variables are reported. The panels are constructed using annual data. Individuals are included in the dataset until at least one of the relevant variables is missing for a given year (i.e., an unbalanced panel). Under this procedure, we are left with 1,168, 1,188, 1,191 and 1,187 individuals in the drug-crime, violent, property and overall crime panels, respectively.[18] Each

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

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

[17]An alternative to our imputation procedure is, at estimation time, to use the probabilities predicted by our model in Section 2.3 to integrate the likelihood for each individual against the distribution of unobserved criminal experience. As a robustness check, we estimated our model using this alternative approach to deal with the unobserved criminal experience. Specifically, for each individual in our dataset, we simulate S possible paths of crime and enrollment decisions from age at first crime to age of entry into the survey, by sampling S draws of the errors in the crime and enrollment equations. For each individual we then calculated S likelihoods, corresponding to each of the S simulated paths. The individual contribution to our overall likelihood is calculated as the average (i.e., the Monte Carlo integral) over these paths. The results were very similar to our baseline estimates, and since this procedure substantially increased the computational burden, we decided not to use it over our simpler imputation procedure.

[18]The sample size in the overall crime sample is not necessarily the largest across all four samples. For instance,

sample includes, at most, eight years for each individual. The attrition rate in the overall crime sample is on average slightly less than 6% per year.[19]

Table 2.1: Pathways to Desistance Descriptive Statistics - Mean and Standard Deviation By Sample

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

Notes:
* Indicates variables that do not vary over time. Summary statistics for these variables are calculated using only the baseline survey.
1. The descriptive statistics reported in this table correspond to data from the combined baseline and follow-up surveys. Each observation is an individual-year pair.
2. The number of observations varies across the four samples since they differ in the number of missing values for each self-reported crime.
3. The crime and enrollment rates reflect the fraction of observations engaged in crime and enrolled in school, respectively.

Table 2.1 reports descriptive statistics for our four samples. There are several statistics that we wish to highlight. First, crime rates in the sample (i.e., the fraction of individual-year pairs in which a crime was committed) are quite high. The violent crime rate is 44%, 29% for property crime, and 21% for drug related crime. These high crime rates (particularly for violent crime) come from the fact that all individuals in the dataset have been convicted of a serious criminal offense at least once, as this is a requirement for entering the dataset. About 14%

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

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

of the sample is female. There is a large percentage of minorities, with blacks and Hispanics representing 40% and 34% of the sample, respectively. Drug use is also quite prominent, with an average of 47%. The average age for the first crime is 10.7 for violent, 11.5 for property, and 13.9 for drug-dealing crimes, illustrating that many of these adolescents start participating in criminal activities well before high school, particularly for violent and property crime.

Table 2.2 reports descriptive statistics for the tests designed to measure cognitive skills. In our empirical analysis we use the two components of IQ separately: the raw WASI Vocabulary Score and the raw WASI Matrix Reasoning Score. However, for interpretability, we report information on the distribution of IQ scores here as well. On average, IQ scores in our sample are substantially below the average score in the general population (100). In fact, almost 90% of individuals have a score below 100. For our measures of cognitive impairment, the Trail-Making scores take one of four values, where the lowest two values indicate either mild/moderate impairment or moderate/severe impairment. In our sample, 21% have some level of cognitive impairment according to Trail-Making A, and 38% under Trail-Making B. The Stroop Test scores take a continuum of values. For each test, scores above 40 are considered "normal". For the Color, Word, and Color/Word tests respectively, 52%, 36%, and 21% had scores below normal.

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

Figures 2.1-2.3 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 2.1 shows how the probability of committing crime depends on the lagged crime decision, and how this evolves with age.[20] Figure 2.2 shows the same for education. There are two important relationships to notice in these figures. First, both crime and education decisions are highly persistent in that individuals who committed crime (enrolled in school) in the previous period are much more likely to commit crime (enroll in school) in the current period. Second, there is some evidence of dynamic selection since, as individuals age, this relationship becomes stronger.

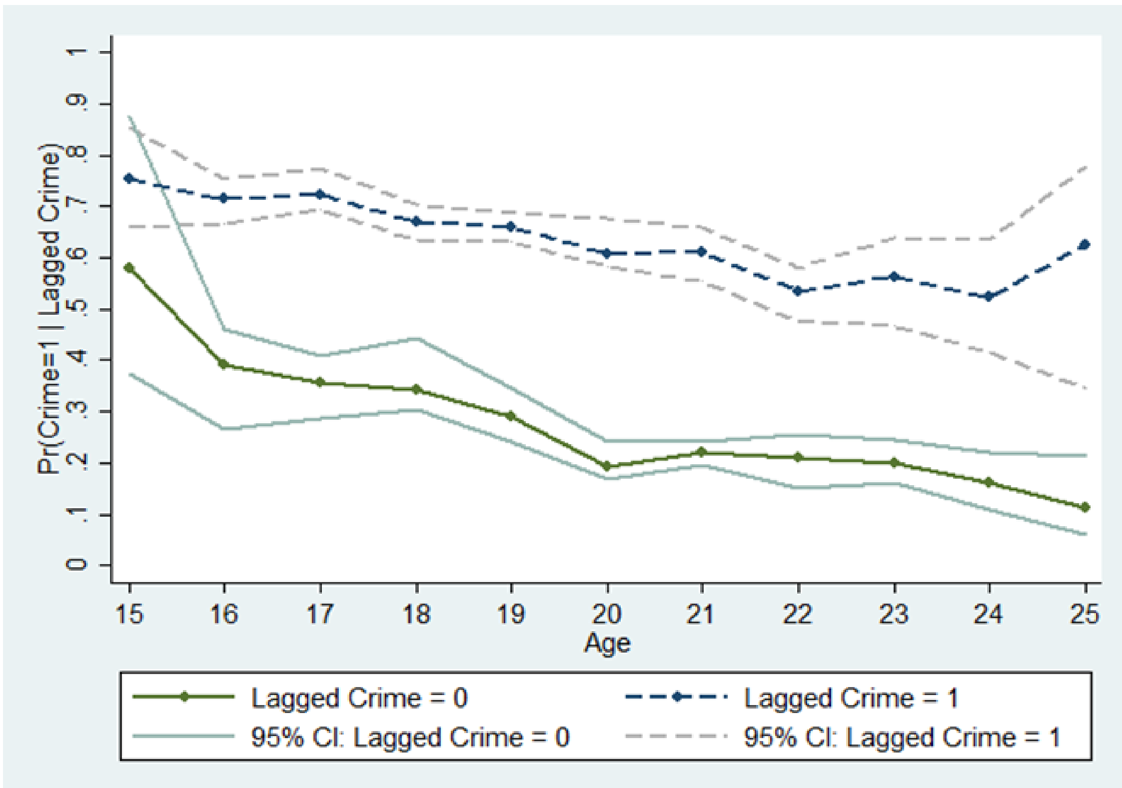
[20]There is a small number of individuals with lagged crime equal to zero at age 15. Since individuals with non-missing values of lagged crime at age 15 entered the sample the previous year, and since committing a crime is what triggers entry into the sample, we would expect all of the people to have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

Table 2.2: 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

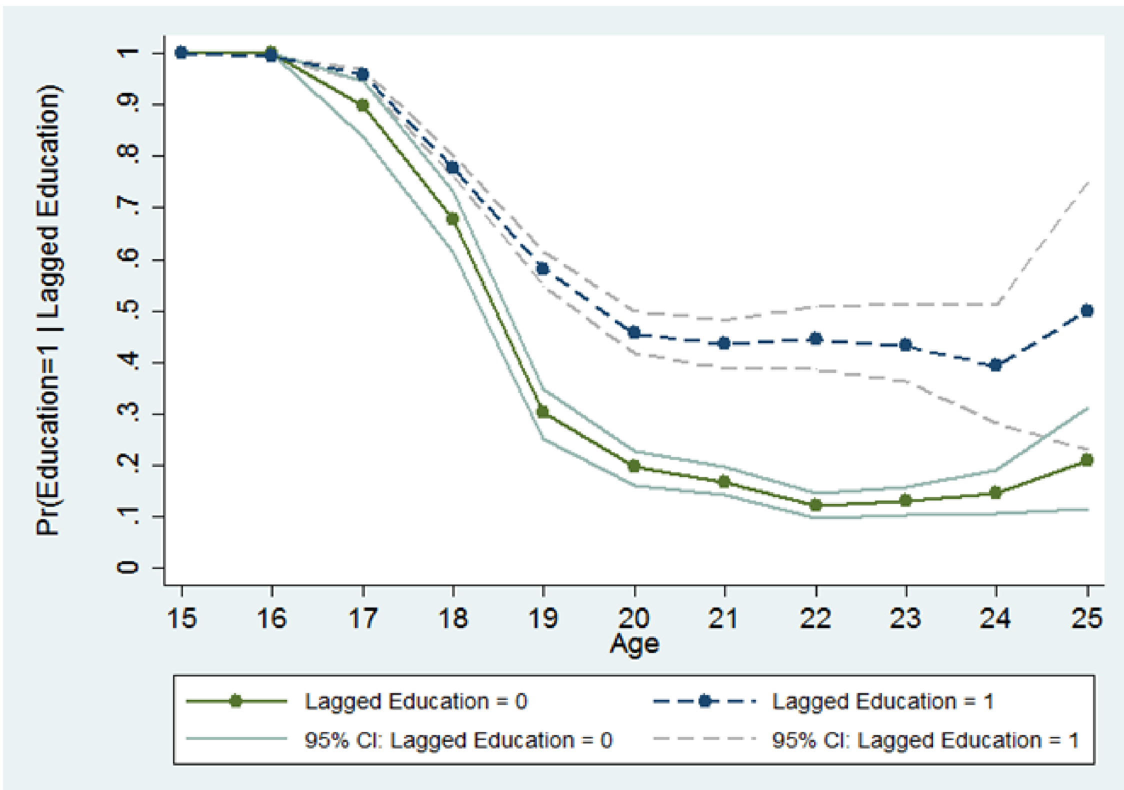
Notes:
1. The descriptive statistics are based on the overall crime sample.
2. The estimate of general intellectual ability (IQ) is based on two subsets: Vocabulary and Matrix Reasoning.
3. The Trail-Making test is a measure of general brain function. Part A involves a series of numbers and the participant 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.

Figure 2.1: Probability of Crime by Lagged Crime Choice and Age



Notes:
1. The figures are based on the overall crime category. For each age category we run a probit of crime on lagged crime. We then predict the probability of engaging in crime by lagged crime and age. The confidence intervals are generated via bootstrapping.
2. Individuals can be considered for the study if they are found guilty of a misdemeanor weapons crime, which we do not categorize into one of our crime types (violent, property, drug). As a consequence, there are a small number of individuals with lagged crime equal to zero at age 15, even though all 15-year-olds entered the survey in the previous year.

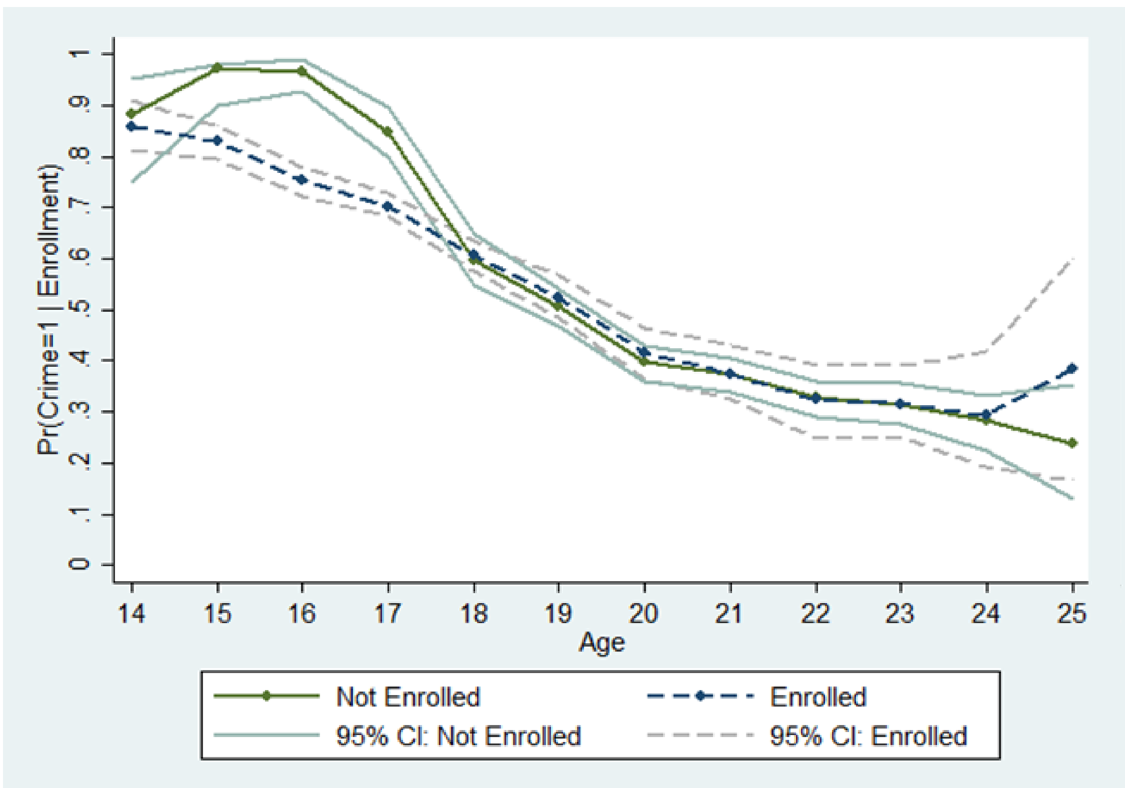
Figure 2.2: Probability of Education by Lagged Education Choice and Age



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

Figures 2.1 and 2.2 demonstrate strong persistence in crime and education decisions. What cannot be determined from the figures alone is the cause of this persistence (Heckman, 1981). This could be generated by persistent differences across individuals that are correlated with education and crime decisions. For example, it may be that low-skill youth are less likely to enroll in school and more likely to commit crimes. A second explanation is that there is state dependence in these decisions. For example, attending school may be easier if the individual has learned the previous year's material. A third possibility is that there are returns to previous experience. It may be the case that individuals become better at committing crimes with more practice, which increases the future probabilities of committing crimes. In our empirical analysis we attempt to disentangle all three potential causes for the observed persistence in decisions.

Figure 2.3: Probability of Crime by Enrollment Status and Age



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

Figure 2.3 illustrates the contemporaneous link between youth crime and enrollment, suggesting a negative correlation in the mid teenage years. While this would seem to suggest a negative effect of 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.

2.3 Model

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

$$v^e_{i,t} = z^e_{i,t}\beta^e + e_{i,t-1}\kappa^e + yc_{i,t}\lambda^e + ye_{i,t}\alpha^e + \eta^e_{i,t}, \tag{2.1}$$

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

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

$$v^c_{i,t} = z^c_{i,t}\beta^c + e_{i,t}\gamma^c + c_{i,t-1}\pi^c + yc_{i,t}\lambda^c + ye_{i,t}\alpha^c + \eta^c_{i,t}, \tag{2.2}$$

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

Notice that in equations (2.1) and (2.2) above, we allow contemporaneous enrollment to affect the crime decision, but not the other way around. The reason for this is that if we were to allow for both types of feedback effects, the resulting model would not be identified due to the problem of incoherency.[21] Therefore, we impose what is referred to in the literature as the coherency condition, by restricting the contemporaneous effect of crime on education to be zero.[22]

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

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

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

[22] We focus on this case because the literature is focused more on the effect of education on crime, as opposed to the effect of crime on education. Alternatively we could assume that the contemporaneous effect of enrollment on crime is zero. In Table A.4 in Appendix A.2, we provide results from the model with the contemporaneous effect in the other direction (crime to enrollment). The results are very similar. In Section 3.5 we discuss how this assumption affects the short-run and long-run impacts on enrollment and crime decisions through simulations of our model.

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

where $\Pr(c_{i,t} = 1 \mid e_{i,t}) = \Pr\left(v^c_{i,t} > 0 \mid e_{i,t}\right)$ and $\Pr(e_{i,t} = 1) = \Pr\left(v^e_{i,t} > 0\right)$, and similarly for the probabilities of $c_{i,t} = 0$ and $e_{i,t} = 0$. If we were to assume that the errors in equations (2.1) and (2.2) are independent and normally distributed, we could estimate the model parameters by estimating separate probits. However, the assumption that the residuals are independent is unlikely to be true, as many of the factors driving enrollment decisions are likely to drive crime decisions as well. When this is the case, $e_{i,t}$ will be endogenous in the crime equation. In order to account for this possibility we use four strategies. First, we include a rich set of individual-level characteristics related to both crime and enrollment decisions, as well as county dummies. Many of these variables (e.g., family crime, certainty of punishment, number of children) are not commonly available, and thus would typically end up included in the error terms.

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

Third, we factor analyze the residuals by taking advantage of some of the unique features of our data. As discussed earlier, one key advantage of our data is that it contains measures of both the cognitive and social/emotional skills of each individual, both of which may be important in driving both enrollment and crime decisions. Using these test measures, we first estimate a correlated factors model to isolate estimates of cognitive and social/emotional skills (see Section 2.3.1 for a description of the factor model we employ). We then include these measures of skills as regressors in our model, by decomposing the errors in equations (2.1) and (2.2 as follows:

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

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

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

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

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

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

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

where the latent variables $v^c_{i,t}$ and $v^e_{i,t}$ are given by

$$
\begin{aligned}
v^e_{i,t} &= z^e_{i,t}\beta^e + e_{i,t-1}\kappa^e + yc_{i,t}\lambda^e + ye_{i,t}\alpha^e + \delta^{e,cog}\bar{\theta}^{cog}_i + \delta^{e,emo}\bar{\theta}^{emo}_i + \varepsilon^e_{i,t}, \\
v^c_{i,t} &= z^c_{i,t}\beta^c + c_{i,t-1}\pi^c + e_{i,t}\gamma^c + yc_{i,t}\lambda^c + ye_{i,t}\alpha^c + \delta^{c,cog}\bar{\theta}^{cog}_i + \delta^{c,emo}\bar{\theta}^{emo}_i + \varepsilon^c_{i,t},
\end{aligned}
\tag{2.3}
$$

and where

$$
\begin{pmatrix} \varepsilon^e_{i,t} \\ \varepsilon^c_{i,t} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right).
$$

2.3.1 Factor Model for Abilities

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

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

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

$$
\begin{aligned}
M^{cog}_{j,i,t_i} &= x_{i,t_i}\beta^{cog}_j + \theta^{cog}_i \delta^{cog}_{j,t_i} + \xi^{cog}_{j,i,t_i}, \\
M^{emo}_{k,i,t} &= x_{i,t}\beta^{emo}_{k,t} + \theta^{emo}_i \delta^{emo}_{k,t} + \xi^{emo}_{k,i,t}.
\end{aligned}
\tag{2.4}
$$

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

$$
\text{If } M^{cog}_{j,i,t_i} = \ell \;\; \Rightarrow \psi_{j,\ell-1} < x_{i,t_i}\beta^{cog}_j + \theta^{cog}_i \delta^{cog}_{j,t_i} + \xi^{cog}_{j,i,t_i} \leq \psi_{j,\ell}. \tag{2.5}
$$

$\theta^{cog}_i, \theta^{emo}_i$ denote cognitive and social/emotional abilities respectively, $\delta^{cog}_{j,t}, \delta^{emo}_{k,t}$ denote loadings that measure the effect of these skills, and the "uniquenesses" $\left\{\xi^{cog}_{j,i,t_i}\right\}_{j=1}^{J}, \left\{\left\{\xi^{emo}_{k,i,t}\right\}_{t=t_i}^{T_i}\right\}_{k=1}^{K}$ capture other determinants of the test scores like measurement error. While we assume that θ_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 Fruehwirth, Navarro, and Takahashi (2016). 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.[25] These are the $\bar{\theta}_i^{cog}, \bar{\theta}_i^{emo}$ we use in equations (2.3).

2.4 Results

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

2.4.1 Factor Analysis

The results from the estimation of the factor model are presented in Tables 2.3-2.4 and Figure 2.4. 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 in the first period WAI Impulse Control measure. Besides being required for identification, these normalizations aid in the interpretation of the factors. Hence, the factor representing cognitive skills is such that an increase of one standard deviation in cognitive skills leads to an increase of one standard deviation in the Matrix Reasoning WASI test, and similarly for the social/emotional factor.

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

In Figure 2.4 we present a variance decomposition that allows us to get an idea of how important it is to account for measurement error (i.e., the uniqueness) when employing these measures. That is, we decompose the variance of the unobservable component of each measurement into the proportion of the variance coming from the skill (i.e., the factor) and the proportion contributed by the uniqueness.[26]

In Tables 2.3 and 2.4 we present the estimated parameters of the factor model for the measurement system. There are two interpretations for the coefficients on the covariates included in the factor model (e.g., gender, race). On the one hand, the coefficients can be interpreted as measuring differences in test scores that are unrelated to skills. For example, under this interpretation, the distribution of skills for men and women is the same, and hence the coefficient on

[25] See Appendix A.1 for details on the estimation of the factor model as well as on prediction.

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

Table 2.3: Estimated Parameters from Factor Analysis - Cognitive Skills

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

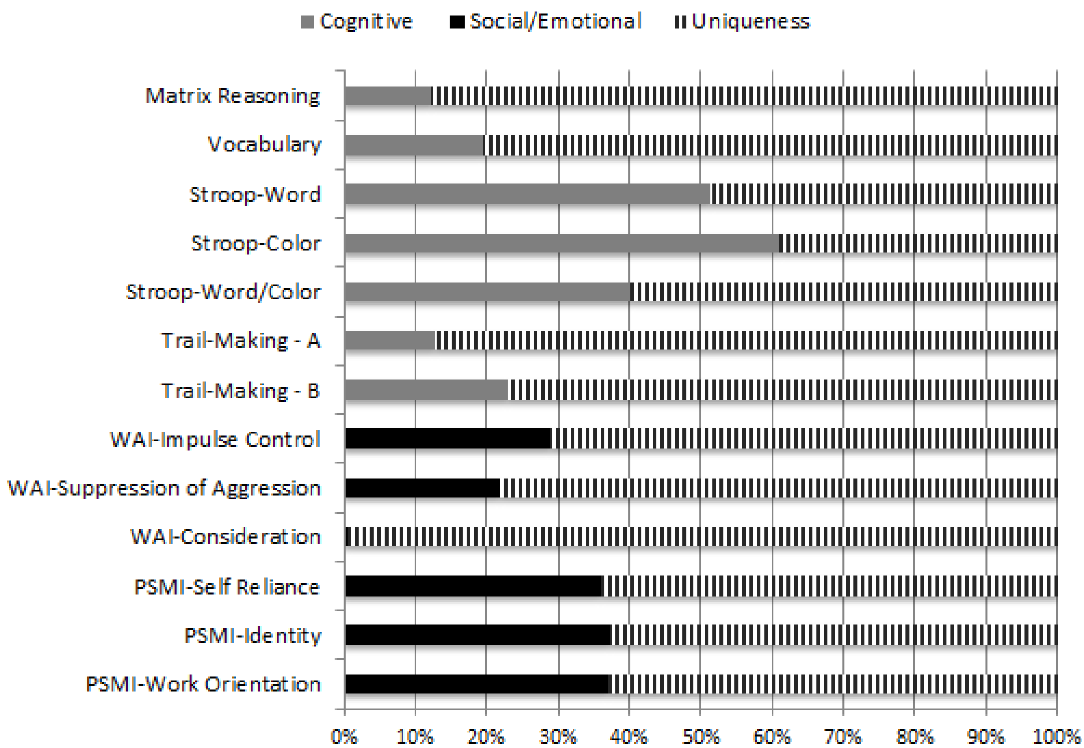
Notes:
1. We estimate a two factor model with cognitive and social/emotional measures. The table presents the parameter estimates for the cognitive measure system. The components of WASI and Stroop are modeled using a linear in parameters specification of the form: $M^{cog}_{j,i,t_i} = x_{i,t_i}\beta^{cog}_j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

Table 2.4: Estimated Parameters from Factor Analysis - Social/Emotional Skills

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

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

Figure 2.4: Fraction of the Variance Explained by Skills



Notes:
1. We estimate a two factor model with cognitive and social/emotional measures. For the cognitive system, the components of WASI and Stroop are modeled using a linear in parameters specification of the form: $M^{cog}_{j,i,t_i} = x_{i,t_i}\beta^{cog}_j + \theta^{cog}_i \delta^{cog}_{j,t_i} + \xi^{cog}_{j,i,t_i}$, where j indexes the measure and i the individual. For the case of Trail-Making we use an ordered model of the form: $M^{cog}_{j,i,t_i} = \ell \Rightarrow \mathbb{1}(\psi_{j,\ell-1} < x_{i,t_i}\beta^{cog}_j + \theta^{cog}_i \delta^{cog}_{j,t_i} + \xi^{cog}_{j,i,t_i} < \psi_{j,\ell})$.
2. For the social/emotional measures we use a linear in parameters specification of the form: $M^{emo}_{k,i,t} = x_{i,t}\beta^{emo}_{k,t} + \theta^{emo}_i \delta^{emo}_{k,t} + \xi^{emo}_{k,i,t}$, where k indexes the measure, i the individual and t age. The figure presents the average fraction of the variance explained by skills. For example, the fraction of the variance of test j explained by cognitive skills is given by: $\frac{1}{T_{cog}} \sum_t \frac{\sigma^2_{\theta,cog}(\delta^{cog}_{j,t})^2}{\sigma^2_{\theta,cog}(\delta^{cog}_{j,t})^2 + \sigma^2_{\xi,cog,j}}$, where T_{cog} is the number of ages for which we observe cognitive scores.

the WASI Matrix Reasoning test of -0.023 in Table 2.3 would be interpreted as indicating that, on average, females perform worse on this test than a male of equivalent skills. On the other hand, they can be viewed as capturing differences in both test-taking and underlying skills.[27] Under this interpretation the coefficient on female reflects a combination of differences in skills and test-taking ability. Without further restrictions we cannot disentangle these two interpretations. Since we also include these variables in the crime and enrollment equations, this implies that our estimates of the coefficients on these variables in the crime and enrollment equations could be interpreted as reflecting combinations of direct effects and indirect effects via differences in skills. It does not, however, affect the interpretation of the other model parameters or of the simulations in Section 2.5.

As can be seen from Table 2.3, having more cognitive skills is related with having "better" scores in all of the cognitive measures we use. The negative sign for the Trail-Making scores is consistent with the way the scores are recorded where a larger score reflects cognitive impairment. As Figure 2.4 shows, our measure of cognitive skills is more related to the Stroop measures of cognitive dysfunction than to the WASI-IQ and Trail-Making measures. However, even for the Stroop measures, cognitive skills can only explain at most 62% of the unobserved variance.

As documented in Table 2.4, for the case of social/emotional scores, more social/emotional skills lead to higher scores for all the social/emotional measures we include. There is also a general pattern consistent with maturation effects, in which the mean scores get better over time (i.e., the constant terms for each period in the equations) and social/emotional skills become a stronger determinant of the scores on the tests (i.e., the loadings). Social/emotional skills explain around 30% of the variance for all measures, except for the WAI-Consideration of Others where it essentially has no explanatory power. This result suggests that our measure of social/emotional skills is more related to individual discipline and control than to attitudes towards other people.

2.4.2 Baseline Model

We now present the results from our baseline specification. In Section 2.4.3, we consider several alternative specifications to evaluate the robustness of our results. In our baseline specification, in order to control for unobserved heterogeneity across individuals, we include our estimated cognitive and social/emotional skill estimates as regressors.[28] The results from the baseline bivariate probit are listed in column 1 of Table 2.5, where we report the average marginal effects of each covariate. We focus on the results for overall crime and discuss the results for the separate crime categories only when the results vary significantly by type of

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

[28] As a robustness check, in Section 2.4.3 we use the set of measurements used to infer the skills as regressors directly.

crime.[29] The results for drug-related, violent, and property crime separately are contained in Tables A.5-A.7 in Appendix A.3.

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

Consistent with what one would expect, having a "non-intact" family, is associated with lower enrollment rates and higher crime rates. Age is negatively associated with enrollment and crime. The result for enrollment is not surprising given that this dataset covers people between the ages of 14 and 26. The finding that crime also decreases with age is consistent with the broader empirical literature on the life-cycle of crime ((Farrington, 1986); (Hirschi and Gottfredson, 1983).[30]

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

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

[30] Drug crime does not seem to decrease with age. Combined with the statistic from Table 2.1 that shows that people start committing drug crimes at much later ages, this suggests that the age profile for drug crime is different compared to violent and property crime ((Sampson and Laub, 2003); (Farrington, 1986); (Wilson and Herrnstein, 1985).

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

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

Table 2.5: Average Marginal Effects from Probits for Crime and Education (Overall Crime)

Variable	**Baseline**		**Controls**		**Uncorrelated Errors**		**No Dynamics**		**Not Instrumenting**		**Cognitive and Social/Emotional Skills**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.049	0.041	0.038	0.030	0.050	0.049	0.051	0.049	0.096	0.044	0.039	0.039
	(0.021)	*(0.020)*	*(0.019)*	*(0.018)*	*(0.021)*	*(0.019)*	*(0.021)*	*(0.021)*	*(0.018)*	*(0.020)*	*(0.022)*	*(0.020)*
Hispanic	-0.025	-0.020			-0.025	-0.022	-0.026	-0.032	-0.026	-0.021	-0.013	-0.024
	(0.015)	*(0.015)*			*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*
Black	0.024	-0.030			0.024	-0.029	0.042	-0.046	0.024	-0.030	0.039	-0.029
	(0.017)	*(0.018)*			*(0.017)*	*(0.018)*	*(0.018)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.018)*	*(0.019)*
Other	0.034	-0.025			0.034	-0.023	0.038	-0.036	0.036	-0.024	0.042	-0.015
	(0.027)	*(0.030)*			*(0.027)*	*(0.030)*	*(0.028)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*
Female	0.058	-0.101	0.054	-0.087	0.058	-0.098	0.070	-0.168	0.057	-0.100	0.053	-0.096
	(0.015)	*(0.016)*	*(0.014)*	*(0.017)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*
Non-intact Family	-0.050	0.031			-0.051	0.028	-0.052	0.040	-0.051	0.030	-0.049	0.031
	(0.015)	*(0.016)*			*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*
Siblings	-0.002	0.004			-0.002	0.003	-0.003	0.006	-0.002	0.004	-0.002	0.004
	(0.002)	*(0.003)*			*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.002)*
Age	-0.080	-0.028	-0.083	-0.035	-0.080	-0.036	-0.104	-0.017	-0.082	-0.031	-0.079	-0.026
	(0.004)	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.004)*	*(0.002)*	*(0.009)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.007)*
Certainty of Punishment	0.003	-0.022			0.003	-0.022	0.005	-0.028	0.003	-0.022	0.002	-0.018
	(0.003)	*(0.003)*			*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.018	0.008			-0.017	0.007	-0.032	0.013	-0.017	0.007	-0.017	0.003
	(0.007)	*(0.007)*			*(0.007)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Family Crime	0.002	0.149			0.002	0.150	-0.002	0.175	0.004	0.149	0.001	0.146
	(0.015)	*(0.015)*			*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*
Drug Use	-0.001	0.224			-0.001	0.225	-0.009	0.267	-0.000	0.224	-0.001	0.204
	(0.012)	*(0.010)*			*(0.012)*	*(0.010)*	*(0.012)*	*(0.012)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.011)*
Unemployment Rate	0.021	0.011	0.021	0.010	0.021	0.014	0.023	0.009	0.037	0.012	0.021	0.010
	(0.006)	*(0.005)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.005)*
Future Outlook Inventory	0.019	-0.024			0.019	-0.023	0.024	-0.030	0.017	-0.024	0.022	0.016
	(0.011)	*(0.011)*			*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.012)*	*(0.013)*
Years of Crime	-0.007	0.020	-0.007	0.039	-0.007	0.020			-0.007	0.020	-0.007	0.017
	(0.003)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.003)*	*(0.003)*			*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Years of Education	0.006	-0.003	0.011	-0.014	0.007	-0.001			0.006	-0.002	0.006	-0.005
	(0.004)	*(0.004)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.004)*			*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Cognitive Factor	0.036	0.014			0.038	0.017	0.041	0.030	0.036	0.015		
	(0.023)	*(0.024)*			*(0.023)*	*(0.024)*	*(0.024)*	*(0.024)*	*(0.023)*	*(0.024)*		
Social/Emotional Factor	0.007	-0.080			0.006	-0.080	0.019	-0.127	0.007	-0.080		
	(0.014)	*(0.014)*			*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*		

Continued on next page...

Table 2.5 – continued from previous page.

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Schools per Young Person	0.322		0.323		0.313		0.319				0.311	
	(0.071)		*(0.072)*		*(0.071)*		*(0.072)*				*(0.071)*	
Lagged Enrollment	0.189		0.191		0.189				0.190		0.185	
	(0.012)		*(0.012)*		*(0.012)*				*(0.012)*		*(0.012)*	
Enrollment		0.088		0.083		0.025		0.202		0.065		0.096
		(0.049)		*(0.053)*		*(0.014)*		*(0.063)*		*(0.051)*		*(0.047)*
Lagged Crime		0.158		0.235		0.159				0.159		0.142
		(0.012)		*(0.013)*		*(0.012)*				*(0.012)*		*(0.012)*
WASI Reasoning Score											-0.002	-0.005
											(0.006)	*(0.007)*
WASI Vocabulary Score											0.000	0.001
											(0.007)	*(0.007)*
Stroop: Color/Word											0.004	-0.010
											(0.007)	*(0.007)*
Stroop: Word											0.009	-0.012
											(0.007)	*(0.008)*
Stroop: Color											-0.003	0.014
											(0.008)	*(0.008)*
Trail-Making: Part B											-0.016	-0.007
											(0.007)	*(0.007)*
Trail-Making: Part A											-0.003	-0.002
											(0.007)	*(0.007)*
WAI - Impulse Response											-0.008	-0.030
											(0.007)	*(0.008)*
WAI - Suppression of Aggression											0.010	-0.044
											(0.007)	*(0.007)*
WAI - Consideration of Others											0.002	-0.027
											(0.006)	*(0.006)*
PSMI - Self Reliance											-0.014	0.023
											(0.010)	*(0.011)*
PSMI - Identity											0.036	-0.018
											(0.010)	*(0.011)*
PSMI - Work Orientation											-0.023	-0.011
											(0.009)	*(0.010)*
Rho	-0.142		-0.137				-0.377		-0.091		-0.157	
	(0.106)		(0.102)				(0.144)		(0.109)		(0.105)	
Observations	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190

Continued on next page...

Table 2.5 – continued from previous page.

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime

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

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

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

As discussed in Section 2.4.1, higher values of our estimates of cognitive and social/emotional skills are associated with better performance on the tests. Therefore, we should expect them to be positively associated with education and negatively associated with crime. We find that higher cognitive skills increase the likelihood of enrollment and higher social/emotional skills lead to lower crime rates. The results imply that a one standard deviation increase in social/emotional skills leads to a decrease in the probability of crime of 3.5%-points. Also, a one standard deviation increase in cognitive skills leads to an increased probability of enrollment of 1.0%-points, although it is not precisely estimated. The effects of cognitive skills on crime and social/emotional skills on education are both small and imprecisely estimated.

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

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

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

For violent and drug-related crime, we find the opposite effect: enrollment leads to an increase in crime rates (10.4%-points for violent and 7.7%-points for drug-related). This suggests the presence of positive complementarities between school and drug/violent crime. This is consistent with the concentration story of Jacob and Lefgren for violent crime–that an increased density of young people leads to more violent interactions. For drug-related crime, one explanation is that the primary buyers of drugs sold by juveniles are other juveniles, and thus attending school allows the sellers of drugs to be closer to their clients.

The last row of Table 2.5 reports the correlation in errors of the crime and enrollment equations. The estimate of -0.142 indicates that the remaining unobserved drivers of crime and education decisions are negatively correlated with each other, although the correlation is not precisely estimated. As we show in the next section, failing to account for this negative correlation leads to a downward bias in the estimate of the contemporaneous effect of enrollment on crime.

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

We find strong evidence of state dependence in both the education and crime decisions ((Brame et al., 2005). Enrolling in school the previous period increases the probability of enrolling in the current period by 18.9%-points. Participating in crime in the previous period increases the probability of crime by 15.8%-points. We also find some evidence of returns to experience, although the effects are smaller. The signs of the results are as expected. An

[33]One potential concern with this result is that not being enrolled is a proxy for being incarcerated, and therefore this estimate captures the incapacitation effect of prison. This is unlikely here, as in our data the relationship between enrollment and incarceration goes in the other direction as they are positively correlated. See also Section 2.4.3 in which we discuss the effects of being in jail in more detail.

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

additional year of education is positively associated with enrollment decisions and negatively associated with crime, but the effects are small and not statistically significant. The effect of criminal experience on crime is positive: an extra year of criminal experience increases the probability of crime by 2.0%-points. The effect on education is negative, with an extra year of crime associated with a decrease in the probability of enrollment by 0.7%-points.

Overall our estimates suggest that there are important dynamics in both the crime and education decisions. While both matter, the effects of state dependence are much larger than the returns to experience. This distinction is relevant for policy, as understanding how the pattern of previous decisions drives current decisions is important for determining how and when to attempt intervention. We discuss this more in Section 2.5 when we illustrate these effects with various simulations based on our model.

The Effect of Education on Crime

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

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

Another difference in our crime measures is that we analyze the extensive margin of crime, whereas aggregate crime measures capture the intensive margin as well. While enrollment may lead to a reduction in the intensive margin, it may not drive it to zero, particularly for our sample of serious offenders, resulting in smaller estimated effects on the extensive margin. In order to examine this, we tried re-estimating our model using continuous measures of crime intensity. The estimated effects of enrollment were negative overall, but small and statistically imprecise, suggesting at most a small role for the intensive margin as an explanation for the

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

differences in our findings. Finally, as discussed in Durlauf, Navarro, and Rivers (2008, 2010), crime regressions based on aggregate data (which is the case for all three papers discussed above) can yield very different results than those based on individual-level data.

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

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

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

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

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

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

In order to explore these alternative explanations, we estimate several additional specifications of our model. We re-estimate our model using a reduced set of controls, specifically only location, non-intact family, age, the unemployment rate, and IQ. We also drop lagged effects and criminal experience and assume that errors across equations are uncorrelated. This roughly corresponds to what the previous literature includes. For both the full and reduced set of controls, we estimate the model on the full sample, the sample of males only (as most of the literature focuses on males), and the sample of males using an alternative measure of years of education for all ages and for those at least 18 years of age. Our alternative measure of years of education does not include years of education obtained while in a locked residential facility. The motivation is that education obtained while incarcerated may have smaller crime-reducing effects. The results are reported in Table 2.6.

Table 2.6: The Effect of Educational Attainment on Crime Alternative Specifications

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

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In the first set of results we include only a reduced set of controls (location, non-intact family, age, unemployment rate, and IQ). In the second set we include the full set of controls from our baseline specification, including lagged decisions, experience, and skills.
3. In columns (3) and (4) we use an alternative measure of years of education that does not include schooling obtained in jail.

As is illustrated in the first set of results, once we drop lags, criminal experience, and the additional controls, our results are closer to those in the literature.[38] The results suggest that an additional year of education is associated with a decrease in the probability of crime ranging from 2.0 to 2.8%-points (4-5%), as we restrict the sample to correspond more closely to what

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

the literature has used. In contrast, the results for the full set of controls are much smaller in magnitude than those in the literature. There is also a small increase in the absolute value of the effect when we employ our alternative measure of education, consistent with quality differences in education obtained while incarcerated.

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

2.4.3 Alternative Specifications

In this section we present results from two sets of alternative specifications to our baseline model that are designed to illustrate how our modeling choices affect the estimates. In columns 2-6 of Table 2.5 we include simple variants to our baseline identification strategy. In particular, we estimate versions of the model in which we incorporate only a limited set of control variables; do not allow for the crime and education equation errors to be correlated (independent probits instead of a bivariate probit); do not allow for dynamics; do not include the number of schools per student as an exclusion in the enrollment equation; and use the direct measures of cognitive and social/emotional skills, as opposed to our estimates of the underlying skills from the factor model.

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

Controls

A key benefit of our data is that we are able to control for a rich set of observable (criminal involvement of the family, expected probability of punishment, and degree of future consideration, among others) and typically unobservable (cognitive and social/emotional skills) sources of individual heterogeneity, that are not commonly available in other datasets. Since most of these variables are highly persistent over time (or fixed), failing to control for them could lead to estimates of the dynamic effects that are biased upwards in absolute value. In order to see the possible extent of this bias, we estimate a version of our model in which we include only a sparse set of individual characteristics and the local unemployment rate. The results are reported in column 2 of Table 2.5. Consistent with our hypothesis, we find that the estimated effects of lagged criminal and educational decisions are inflated, particularly their effects on

[39] We present these results in Tables A.1-A.4 in Appendix A.2 and in Tables A.8-A.13 in Appendix A.3 for the crime-specific estimates.

crime. The returns to criminal experience on crime almost double from 2.0 to 3.9%-points, and the effect of lagged crime increases by roughly 50% from 15.8 to 23.5%-points. The effects of educational experience on both crime and enrollment also increase and become statistically significant (from -0.3 to -1.4%-points and from 0.6 to 1.1%-points, respectively).

Uncorrelated Errors

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

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

Not Instrumenting

As we discuss above in Section 2.3, in order to address the potential endogeneity of enrollment in the crime equation, we introduce an exclusion restriction by adding the change in the number of schools per person in the enrollment equation. In column 5 of Table 2.5 we present results in which we do not include this, in order to illustrate its effect on our estimates. The primary concern was that failing to appropriately control for endogeneity would lead to a biased estimate of the effect of enrollment on crime, which could in turn generate bias in the other estimates as well. We find that by not including this variable, the estimate for contemporaneous enrollment drops from 8.8 to 6.5%-points. The difference is consistent with the expected bias given the negative correlation of the errors. This result demonstrates that there is some bias that this exclusion restriction is correcting for. However, the bias is not particularly large, which is likely due to the fact that our data allow us to control for many sources of observed and unobserved heterogeneity that would otherwise generate further correlation in the errors of the crime and enrollment decisions, and exacerbate the endogeneity problem.

Cognitive and Social/Emotional Skills

We also estimate a specification in which we replace our estimates of skills with the measures used to infer them. This allows us to investigate whether our results are sensitive to our use of the estimated cognitive and social/emotional skills, and also to better understand how cognitive and social/emotional skills contribute to enrollment and crime decisions. As can be seen in column 6 of Table 2.5, the estimates on the other variables are very similar to the baseline estimates, illustrating that our factor-model-generated measures are effective summaries of these skills.

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

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

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

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

Modeling Choices While in Jail

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

In order to determine if our results are sensitive to this choice,[40] we estimate three other model specifications. In the first, we set enrollment to zero if an individual did not attend a community school (i.e., attended an institutional school only, or attended no school). In the second specification, we add a variable to the model that is an indicator for whether the individual was incarcerated at the time of the interview, to allow for being in jail to affect choices. Finally, we add the indicator for jail interacted with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail. The results of the three specifications are reported in Table A.1 in Appendix A.2.

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

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

here is that it is more difficult to obtain and use drugs while in jail, so adding these observations (in jail and not attending school) generates a positive correlation between drug use and enrollment.

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

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

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

Drug Use

Another potential concern relates to the fact that drug use is a choice rather than an exogenous variable, which may bias some of our results. In particular one might think that education affects the propensity to use drugs, and that our finding that drug use has a strong positive effect on crime, and educational attainment does not, masks the indirect effect of education on crime via drug use. In order to check for this possibility, we estimated a specification of our model in which we drop drug use. The results are reported in column 1 of Table A.2 in Appendix A.2. Dropping drug use does not change the effect of education on crime through educational attainment, suggesting that education does not have much effect on crime either directly or indirectly through drug use. On the other hand, dropping drug use increases the estimates of the effects of both skill measures on crime, by about 4%-points each. This suggests that skills have not only a direct effect (which is what we capture in the baseline estimates), but also an indirect effect through drug use. The results for the other coefficients are largely unchanged.

Defining Enrollment

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

Age-Varying Coefficients

One potential concern with our baseline specification is that, if the effects of previous and contemporaneous education and crime decisions vary by age, then any estimated effects, particularly long-run effects, may be biased. In order to examine whether, and to what extent, this may be the case, we estimate a version of the model in which we allow the effects of accumulated experience, lagged decisions, and contemporaneous enrollment to vary by the age of the individual. In particular, we interact these variables with a dummy for whether the person is over 19 years old. In column 1 of Table A.3 in Appendix A.2 we find that the estimates vary slightly by age, but the differences are small. The largest change is in the effect of lagged enrollment on education, in which the marginal effect decreases with age from 22.9 to 17.0%-points, suggesting that the state dependence in educational decisions decreases slightly as individuals age, which is not surprising. Overall, the results seem to be consistent across age.

Criminal Experience

In the baseline survey we observe the age at which individuals first engage in crime, but we do not have a measure of accumulated criminal experience at the time of entry into the survey. In our baseline model we impute the accumulated years of crime using the procedure described in Section 2.2. Our estimates suggest a larger role for state dependence compared to returns to experience. One possible explanation for this result is that experience enters utility in a non-linear fashion, causing us to not fully capture its impact, whereas lagged crime is a dummy variable, and therefore already enters the model flexibly.

In columns 2 and 3 of Table A.3 in Appendix A.2, we allow experience to enter quadratically and as a piecewise-linear function of experience, allowing for different returns for 0-4, 5-9, and 10+ years of criminal experience. When we allow experience to enter quadratically we find that, consistent with the baseline, criminal experience has a small negative effect on

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

enrollment and a positive and increasing effect on crime. However, we lose statistical significance on all of the associated parameters. In the second specification, the effect of years of crime on crime is similar to the baseline with no significant variation across the different experience categories.

Another concern is that our imputation procedure generates a noisy measure of criminal experience, making it more difficult to tease out the true returns to experience. In column 4 of Table A.3 in Appendix A.2, we use only the observed accumulated experience after entry, interacted with age of entry dummies, instead of our imputed measure.[42] Since observed criminal experience is likely to be positively correlated with the unobserved experience that occurs prior to entering the survey, we should expect that the coefficients will be inflated, as they will capture the effect of both the observed and unobserved experience. This upward bias in the coefficients is likely to be increasing in the age of entry into the survey, since the unobserved period is longer for people who entered the survey at an older age. This is consistent with our estimates. Furthermore, even if we ignore the bias in these coefficients, we still find a larger impact of lagged crime compared to criminal experience. Overall we conclude that our finding that state dependence has a stronger effect on crime than criminal experience is not driven by measurement or specification issues related to experience.

The Contemporaneous Effect of Crime on Education

In our baseline model, we estimate the contemporaneous effect of education on crime. As discussed above, we could have alternatively estimated the contemporaneous effect from crime to education.[43] In Table A.4 in Appendix A.2 we present estimates from this alternative specification. The results in the first column show that contemporaneous crime leads to an increase in enrollment of 9.7%-points, which is similar in magnitude to our estimate of the effect of enrollment on crime in our baseline specification. The correlation in errors of the crime and enrollment equations is negative, although not precisely estimated, as was the case in the baseline. The results for other coefficients are relatively unchanged. In column 2 we include lagged state arrest rates as an exclusion restriction in the crime equation (to serve as an instrument for contemporaneous crime in the enrollment equation).[44] The coefficient on contemporaneous crime increases slightly to about 12%-points, while the other coefficients remain largely unaffected.

2.5 Model Simulations

In this section we attempt to disentangle the roles of state dependence (i.e., lagged choices), criminal and human capital (i.e., accumulated years of crime and education), and heterogeneity

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

[43]But not both. See our discussion in Section 2.3.

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

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

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

2.5.1 Dynamic Effects of Temporary Differences

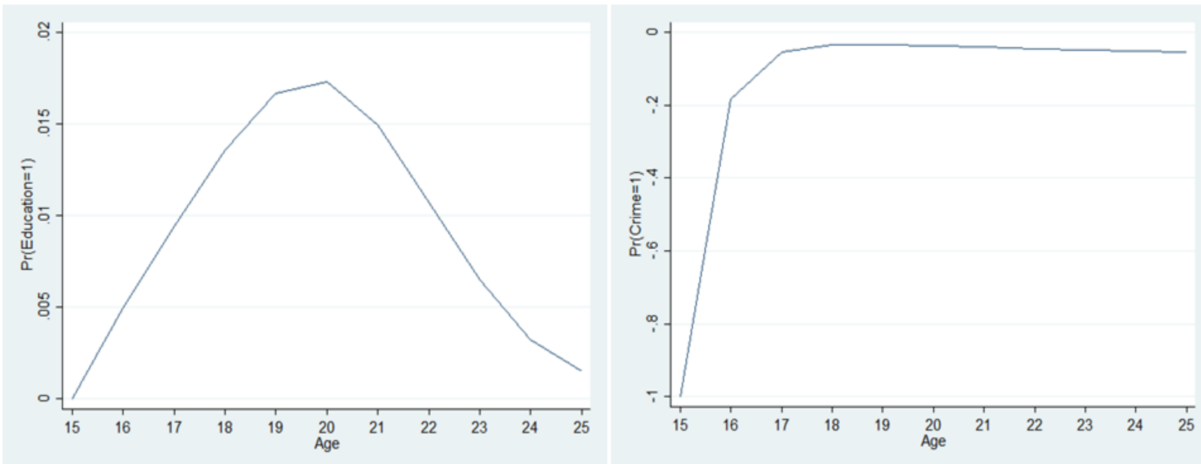
We begin by simulating the differences from committing versus not committing a crime at age 15. Figure 2.5 shows that this has a very small effect on the probability of enrolling over time. The probability differs by 1.7%-points after 5 periods (from a baseline of 40%), and then it decreases as a consequence of aging since, after 10 years, almost no one in the data is enrolled anymore. Figure 2.5 shows that the effects on crime are much larger. Mechanically, the difference in the probability of committing a crime at age 15 is one. After one year, the probability of committing a crime is lower by 20%-points, from a baseline of 70%. This effect is almost entirely a consequence of state dependence (i.e., lagged crime). After that, the effect

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

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

diminishes over time but, because of the decrease in criminal experience, it does not disappear. After 10 years, the person who did not commit a crime at age 15 is approximately 6%-points less likely to commit a crime.

Figure 2.5: No Crime at Age 15 - Effect on Average Probability of Education and Crime

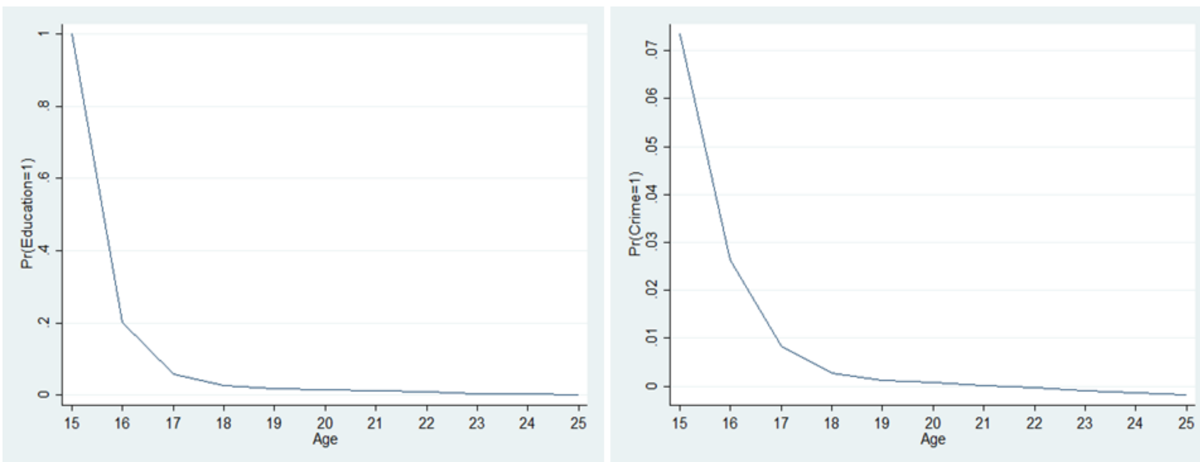


Notes:
1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are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

Next, we analyze enrollment in school at age 15. In Figure 2.6 we can see that the effect of education on education is very similar to the effect of crime on crime. Mechanically the difference in the probability of being enrolled is one at age 15. As a consequence of state dependence, the probability is around 20%-points higher after a year. It decreases over time, reaching zero after 10 years. Its effect on crime is small but not insignificant (at least in the first years). Since enrollment has a positive contemporaneous effect on crime, as we can see in Figure 2.6, it increases the probability of crime by 8%-points initially. The effect rapidly decreases, and it reaches zero after 3 years. After that, it becomes slightly negative but very small as more and more human capital (i.e., years of education) gets accumulated.[47]

[47] As we mention in Section 2.3, we also estimated a version of the model in which there is a contemporaneous effect of crime on enrollment instead of an effect of enrollment on crime. In all of the simulations we describe in this section the long-run outcomes are very similar between the two model specifications. In a few cases, the short-run effects are different. In particular, for the case of the difference in committing a crime at age 15, in the alternative specification there is a short-run negative effect on enrollment that does not appear in our baseline model. Similarly, for the case of a difference in enrolling in school at age 15, there is no longer a short-run positive difference in crime. In both cases these differences diminish quickly. Figures for simulations from this alternative specification that are analogues to Figures 2.5-2.10 are located in Figures A.1-A.6 in Appendix A.2.

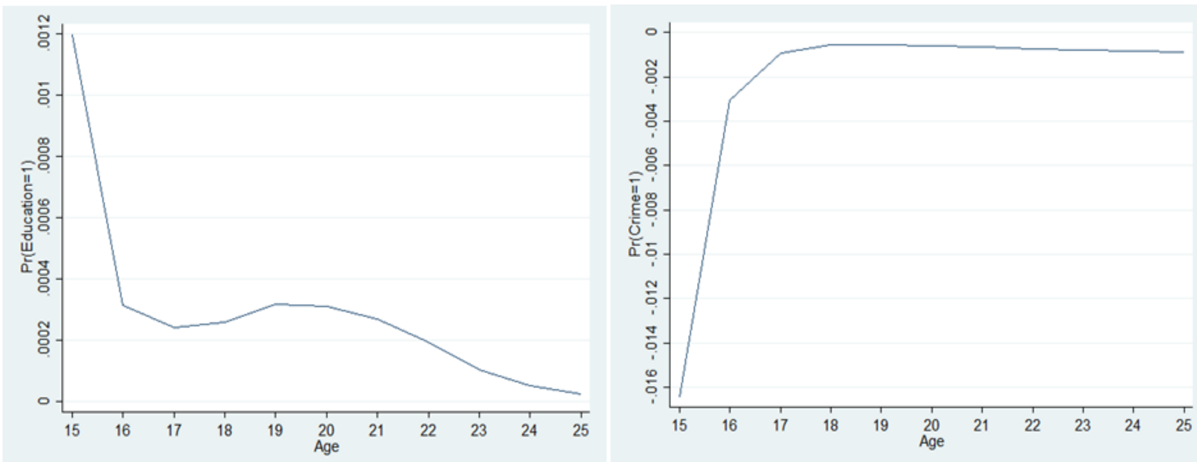
Figure 2.6: Enrolled at Age 15 - Effect on Average Probability of Education and Crime



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

The third simulation we present, the effect associated with a 10%-point difference in the perceived probability of punishment at age 15, is shown in Figure 2.7. The effect on enrollment is negligible. Its effect on crime, on the other hand, is larger. At age 15, it reduces the probability of committing a crime by almost 2%-points. While the effect decreases rapidly, 10 years later there is a 0.1%-point lower probability of committing a crime.

Figure 2.7: Increase in Certainty of Punishment at Age 15 - Effect on Average Probability of Education and Crime



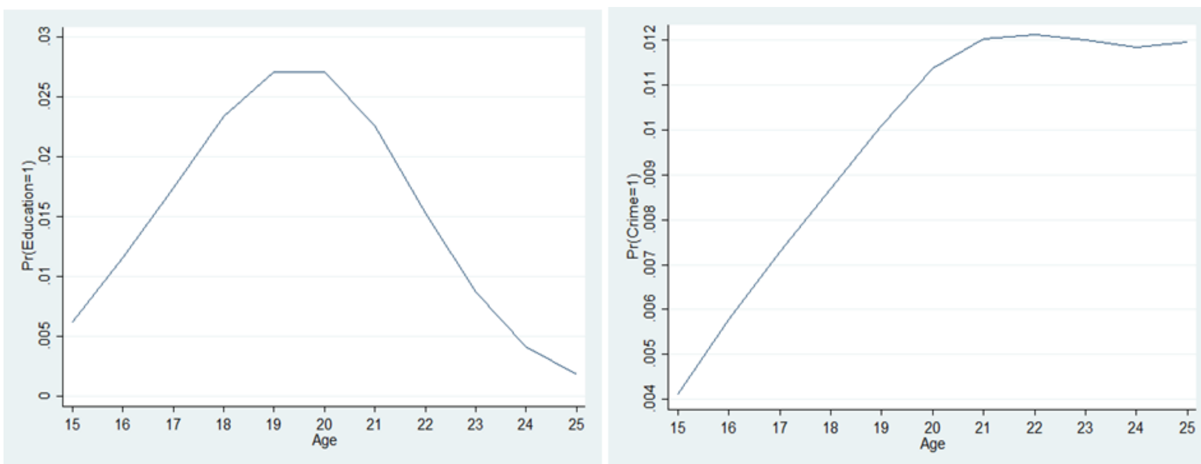
Notes:
1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are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

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

2.5.2 Dynamic Effects of Permanent Differences

We next consider the effects that permanent differences in heterogeneity (while holding all other characteristics at their median values) may have on both the enrollment and crime probabilities. We begin by simulating paths of an individual with cognitive skills at the 25th percentile and comparing to one with skills at the 75th percentile in the data. While this may sound like a large difference, this is for individuals in our selected data where this distribution is much more compressed than in the overall population. For example, the 25th and 75th percentiles of the cognitive skill distribution are associated with IQ scores of 89 and 98 and scores of 39 and 48 on the Stroop Word test, respectively–a modest difference.[48] Figure 2.8 shows the effect on enrollment. Not surprisingly, higher cognitive skills are associated with a larger probability of being enrolled, but the magnitude of the difference is small: at most 3%-points (after five years). Cognitive skills are essentially not related to the probability of crime.

Figure 2.8: Cognitive Factor 25th versus 75th Percentile - Effect on Average Probability of Education and Crime



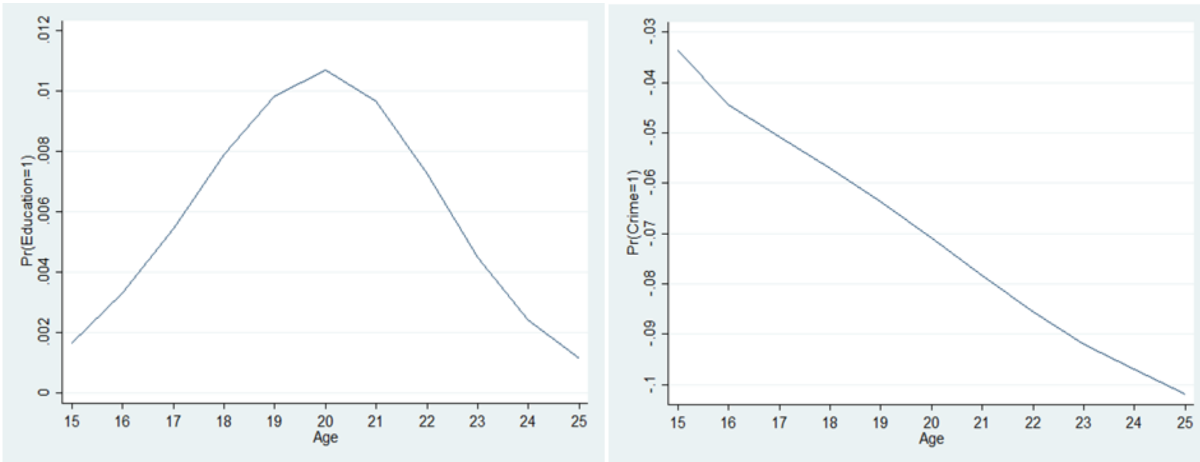
Notes:
1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are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

Figure 2.9 shows similar results for social/emotional skills. A movement from the 25th to 75th percentile for these skills is equivalent to a one-third of a standard deviation difference in impulse control, for example. As can be seen from the figures, the effect on enrollment is negligible. A different story arises when we look at the effect on criminal activity. The probability of committing a crime is lower by 3%-points for the individual with higher social/emotional skills at age 15, and the effect keeps growing over time. After 10 years the probability of committing a crime is reduced by 10%-points.

The final simulation is shown in Figure 2.10. In this case we simulate the paths based on a permanent 10%-point difference in the perceived probability of punishment. After five years the probability of enrollment is marginally larger, by less than 0.7%-points. The impact on crime is more significant. At age 15, the probability of crime is almost 2%-points lower for the

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

Figure 2.9: Social/Emotional Factor 25th versus 75th Percentile - Effect on Average Probability of Education and Crime



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

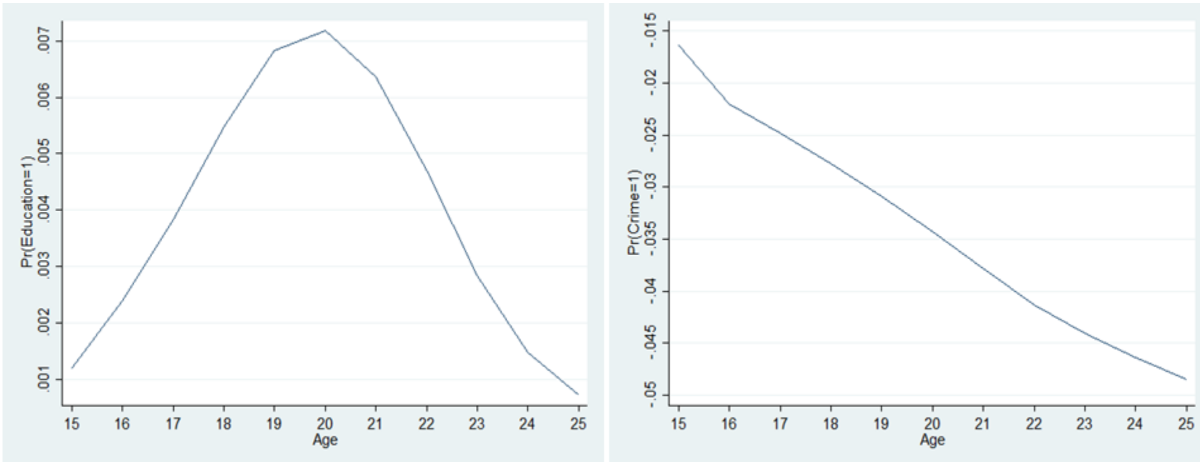
individual with the higher perceived probability of punishment and the difference gets larger over time. After ten years it is almost 5%-points.

2.6 Conclusion

In this chapter, 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 behaviour, and for the purpose of designing policy. We find that individual heterogeneity is strongly related to criminal behaviour. 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 behaviour. In particular, we find that social/emotional skills are important drivers of criminal behaviour.

While we do not directly simulate potential policies designed to increase enrollment and/or decrease crime, our model simulations illustrate how policies targeted at altering individual heterogeneity (e.g., social/emotional skills) would drive changes in education and crime over time. We find, perhaps unsurprisingly, that permanent or long-run changes generate the largest

Figure 2.10: Increase in Certainty of Punishment (Permanent) - Effect on Average Probability of Education and Crime



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

effects. However, policies with temporary changes to individual behaviour, such as keeping people out of crime for one period, can also have lasting effects. For example, a policy that prevents someone from committing a crime in a given year would generate an effect on crime in the following year of -18%-points. This implies that there is room for policies designed to shock individuals out of current bad decisions, and thus break the persistence caused by this state dependence. To the extent that these types of policies are easier to implement than permanent changes to individuals, their effect should not be dismissed. The reductions obtained are considerable and, at least in the case we model here, they are obtained during the ages in which criminal activities are at their peak.

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

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

Bibliography

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

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

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

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

Betts, Julian R. and Laurel L. McFarland. 1995. "Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments." *Journal of Human Resources* 30 (4):741–765.

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

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

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman. 2003. "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice." *International Economic Review* 44 (2):361–422. 2001 Lawrence R. Klein Lecture.

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

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

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

Chiteji, Ngina. 2010. "Time Preference, Noncognitive Skills and Well Being Across the Life Course: Do Noncognitive Skills Encourage Healthy Behavior?" *American Economic Review* 100 (2):200–204.

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

Cooley Fruehwirth, Jane, Salvador Navarro, and Yuya Takahashi. 2016. "Identification and Estimation of Time-varying Treatment Effects: How The Timing of Grade Retention Affects Outcomes." *Journal of Labor Economics* 34 (4). Forthcoming.

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

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

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

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

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

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

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

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

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

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

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

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

Grogger, Jeffrey. 1995. "The Effect of Arrests on the Employment and Earnings of Young Men." *The Quarterly Journal of Economics* 110 (1):51–71.

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

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

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

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

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

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

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

Hirschi, Travis and Michael Gottfredson. 1983. "Age and the Explanation of Crime." *American Journal of Sociology* 89 (3):552–584.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Luallen, Jeremy. 2006. "School's Out...Forever: A Study of Juvenile Crime, At-Risk Youths and Teacher Strikes." *Journal of Urban Economics* 59 (1):75–103.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Piquero, Alex R, Carol A Schubert, and Robert Brame. 2014. "Comparing Official and Self-report Records of Offending across Gender and Race/Ethnicity in a Longitudinal Study of Serious Youthful Offenders." *Journal of Research in Crime and Delinquency* 51 (4):526–555.

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

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

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

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

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

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

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

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

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

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

Chapter 3

The Labour Market for Disadvantaged Young Individuals

3.1 Introduction

The school to work transition and the early career path can be a difficult one for young people, especially among the low-educated. These disadvantaged individuals tend to have low employment rates, experience long periods of unemployment before their first job, and face lower wages than their older counterparts (Wolpin, 1987; Eckstein and Wolpin, 1995; Bowlus, Kiefer, and Neumann, 2001). A large fraction of these youth are also engaged in crime. The literature has documented that crime is widespread among young low-educated males in poor urban areas (Freeman, 2000; Raphael and Sills, 2007). Many make money from illegal activities, and some even have a career in crime, spending most of their time engaged in criminal activities. For instance, around 30% of low-educated young males in the United States reported an income from crime in 1980 (Lochner, 2004). Youth also account for a large share of total crime. In particular, 1.9 million youth between the ages of 15 and 19 were arrested in 2010, accounting for 19% of all arrests, despite representing only 7% of the total population.[1]

In this chapter, I characterize the criminal and legal labour sectors for disadvantaged young individuals and empirically investigate the factors driving the transitions between sectors. I focus on the role of heterogeneity, earnings, human capital and criminal capital accumulation in determining transitions across the criminal and legal labour sectors. Analyzing what the labour market looks like for disadvantaged young individuals, and how the different alternatives interact with each other (i.e., crime and legal employment), is key to interpreting how individuals make employment choices. Understanding the mobility patterns across sectors can provide guidelines for future research aimed at explaining crime and/or legal employment choice. To the extent that the criminal and legal labour sectors interact with each other, criminal behaviour should not be studied in isolation of legal employment behaviour.

[1] This figure is based on data from the FBI's Uniform Crime Reports.

The data I employ comes from the Pathways to Desistance Study (PDS), 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 employment in the legal sector, educational decisions and other outcomes. As a result, the dataset contains a rich panel of information about decisions to participate in crime and the legal labour sector. 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.

The PDS is especially well-suited for understanding the labour market for disadvantaged young individuals. First, it contains a large share of minorities and low-educated individuals. It also concentrates on young offenders; a group that contributes significantly to aggregate crime rates and, at the same time, faces important challenges in the legal labour sector. Furthermore, besides containing comprehensive information regarding legal employment the PDS also has very detailed information regarding participation in the criminal sector including criminal earnings, types of crimes, and the number of weeks participating in crime in a given month, which allows me to characterize both employment sectors.

This chapter is related to the literature that analyzes the legal labour sector for disadvantaged young individuals (Freeman and Wise, 1982; Freeman, 1991), and the challenges faced by individuals with criminal histories (Bushway, 2004; Bushway and Sweeten, 2007; Raphael, 2011). It is also related to some recent work that investigates the implications of the legal labour sector on the criminal sector, and viceversa. Several papers have found a strong association between the inability to obtain legal employment and criminal activity (Raphael and Winter-Ebmer, 2001; Gould, Weinberg, and Mustard, 2002; Lochner, 2004; Schnepel, 2016), as well as low wages and crime (Grogger, 1998; Machin and Meghir, 2004; Kling, 2006). Unlike most of these studies, this chapter concentrates on a group of individuals who is rarely studied and who mostly contributes to crime participation. Furthermore, I not only characterize the main features of the legal labour sector faced by these young individuals, but I do a similar analysis on the criminal sector. The extensive data on criminal activity and legal labour sector participation available in the PDS also allows me to develop a good understanding of the interactions between the criminal and the legal labour sector.

As a preview of the main results, I find that disadvantaged young individuals face two low-quality employment alternatives. On the one hand, jobs in the legal labour sector are characterized by short duration and low wages. Consistent with their low quality, these jobs present small returns to legal labour market experience. Income crime activity presents similar features as legal jobs: short average duration and small returns to experience. The criminal sector offers an earnings premium relative to the legal labour sector, which partially compensates for the inherent risk of the activity. I find that the transitions between these two sectors are related to cognitive and social/emotional skills and experience in the corresponding sector. This last fact suggest that even if experience has a small role for earnings it is possibly related to other aspects of the activities (e.g., availability of opportunities). Lastly, I provide evidence that earnings in the criminal and legal labour sectors play a significant role on the transitions across sectors. These results imply that crime and employment choices among disadvantaged young individuals are strongly related and should not be studied in isolation of each other.

The rest of the document is organized as follows. Section 3.2 adds more relevant data details on top of Section 2.2 in Chapter 2. Section 3.3 characterizes the labour market for disadvantaged individuals and Section 3.4 analyzes the main patterns of mobility between the criminal and legal labour sectors. Finally, Section 3.5 concludes.

3.2 Data

I use data from the Pathways to Desistance Study (PDS), which was described in detail in Chapter 2. In this chapter, however, I use the Calendar Data rather than the Annual Data introduced in Chapter 2.[2] The Calendar Data is collected at a monthly level across distinct domains including academic achievement, criminal activity, legal employment, contacts with the justice system, among others. The monthly data allow me to construct criminal and legal employment histories. As a result, this type of data is well-suited for characterizing the criminal and legal labour sectors, as well as understanding the dynamics across sectors.

Regarding legal employment, the calendars collect information on the number of jobs an individual holds in a given month, including part-time and under-the-table jobs. For each job, the survey gathers information on the number of hours worked, number of weeks worked, hourly wage, and type of job. The survey also keeps track of each job from month-to-month.

The monthly data on criminal activity come from self-reporting by each respondent. In order to encourage accurate self-reporting, responses are kept confidential and participants were given a certificate of confidentiality from the U.S. Department of Justice. The self-reported offenses consist of 24 components, each of which relates to involvement in a different type of crime, and which can be used to identify whether the adolescent commits a criminal act in a given month.[3] In addition, the survey gathers information on weekly earnings from selling drugs, selling stolen property, stealing merchandise, gambling and prostitution, as well as the number of weeks worked in these activities in a given month.

The survey also indicates whether individuals were locked in a facility in a given month, along with the amount of days spent in each facility and the type of institution (e.g., jail, prison, detention center, Pennsylvania Youth Development Centers, and the Arizona Department of Juvenile Corrections).

Altogether, I obtain a monthly panel of individuals making choices across the criminal and legal labour sectors. For the analysis in this chapter, I focus on males who are no longer at-

[2]While the Annual Data is accessible without restriction, the Calendar Data requires users to request access to the data and meet specific requirements.

[3]The 24 self reported offenses are: destroy property, set fire, enter a building to steal, shoplift, buy, sell or receive stolen property, use credit card illegally, steal a car or motorcycle, sell marijuana or other illegal drugs, carjack someone, drive drunk or high, pay for sex, force sex upon another person, kill someone, shoot someone, rob someone with a weapon, beat up someone, engage in a fight, carry a gun, enter a car to steal, and go joyriding.

tending school, and therefore, are only deciding between unemployment, legal employment, and crime. As a consequence, I concentrate on a sample of males who do not pursue further schooling beyond age 19 and explore their behaviour once they have transitioned out of school. I choose the age threshold of 19 to exclude individuals who are more educated, and that potentially face a different labour market. Moreover, males older than 19, who only obtain schooling if they are incarcerated or who are enrolled in GED programs, are not excluded from the sample. Schooling obtained in prison or through GED programs should not significantly boost the chances of getting a legal job or any other labour market outcome relative to the rest of the sample (Cameron and Heckman, 1993).

For the analysis, a legal job is defined as an employment relationship in the legal labour sector that consists of at least 20 hours per week, including under-the-table jobs. At each point in time, individuals in the data can hold multiple jobs in the legal labour sector. Although this is rarely seen in the sample, I follow the literature to deal with overlapping jobs. Those legal jobs that are completely covered by another job are dropped. For legal jobs that partly overlap, the starting date of the second job is replaced by the stopping date of the first job. For legal jobs that completely overlap, the job with the higher wage is used. I am then left with the main source of earnings in the legal labour sector for each individual in each month. To deal with outliers in monthly legal earnings, earnings in the data are trimmed based on the 1st and 99th percentile from the Current Population Survey for high school dropouts in the same age range as in the PDS, for each year in the sample (i.e., 2001 to 2009).

In the criminal sector I focus on income crime, which encompasses any illegal activity aimed at earning money.[4] Income crime includes selling stolen property, selling drugs, stealing merchandise, gambling, and prostitution. An income crime spell starts when the individual participates in income crime for at least two weeks in a given month. The spell ends when the individual participates in income crime for less than two weeks in a given month or if he does not participate at all. To deal with outliers in monthly criminal earnings, these are trimmed 1% at the top and bottom of the distribution.[5] I deflate earnings in both sectors to the 2000 level using the US Bureau of labour Statistics CPI.

The final sample consists of 527 males.[6] Individuals are included in the sample until at least one of the relevant variables is missing for a given month.[7] Table 3.1 reports descriptive

[4]In Section 3.4, I explore the role of non-income crime on transitions to income crime, legal employment, and incarceration.

[5]Different from the legal labour sector, there is no external study, such as the CPS, that can be used to obtain reliable bands for criminal earnings.

[6]The original sample of monthly data starts with 1,265 individuals that complete the first follow-up survey after the baseline survey. Females are excluded from the sample. Of the 1094 males left, 554 are still going to school after age 19, and 13 have missing data on participation in the legal labour sector, criminal sector, or incarceration that does not allow me to determine the state.

[7]The attrition rate in the sample is on average slightly less than 5.5% per year. One concern is that individuals who leave the sample early are different from individuals who stay until the end of the survey. As a robustness check, I estimated a probit model for attrition in the sample and found that it is not related to race, cognitive and social/emotional skills, education, self-reported criminal activity, the age at which the individual was first arrested, or the average number of arrests per year prior to the baseline survey. There is, however, some evidence of selection based on the average number of arrests per year after the baseline survey, although the effect is small.

statistics for the sample. There are several statistics that I wish to highlight. First, there is a large percentage of minorities and low-educated individuals. Blacks and Hispanics represent 37.8% and 37.6% of the sample, respectively. The share of individuals with a high school degree is 24.5%.[8] Not surprisingly, the crime rate in the sample (i.e., the fraction of individual-month pairs in which an income crime is committed) is quite high. The monthly crime rate is 10.6%, with an average annual crime rate of 29.8%. The monthly employment rate is 34.4%, and 71.5% of the individuals hold a legal job at some point in the sample. Hence, even if the sample contains disadvantaged young individuals who were found guilty of a serious crime, most of them do not seem to be fully banned from working in the legal labour sector.[9] Lastly, there is overlap between sectors.[10] Around 33.9% of the time, participation in income crime is accompanied by simultaneous participation in the legal labour sector in the same month.

3.3 Labour Market Facts: Criminal and Legal Labour Sectors

This section aims at characterizing both the criminal and legal labour sectors for disadvantaged young individuals using the data described in Section 3.2. I start by analyzing the main features of the legal labour sector, including types of jobs, average duration, and earnings. I then present a descriptive picture of the criminal sector and compare it to the legal labour sector. Understanding what the criminal sector and the legal labour sectors look like for disadvantaged young individuals is key to interpreting how they make unemployment, legal employment, and crime choices.

Table 3.2 displays the cross-section distribution of legally employed individuals across occupations. Manual and skilled occupations, such as cutting grass and carpentry, represent around 62.3% of the jobs. Not surprisingly, a large share of individuals take restaurant jobs (15.4%). Lastly, less than 15% have clerical, managerial, or administrative positions. Overall, most of the jobs are low-skilled. This is not surprising since a large portion of employment opportunities are not accessible for disadvantaged individuals, for example, due to education requirements. Beyond education, there is evidence that employer hiring preferences may be further affecting the types of job available to individuals with criminal records (Agan and Starr, 2017; Pager, 2003; Uggen et al., 2014). Firms that restrain from hiring individuals with criminal records tend to be filling positions with more educated people relative to firms that are willing to hire individuals with criminal past (Raphael, 2011). Furthermore, criminal records

These results are available upon request.

[8]Relative to the youth population of Philadelphia and Phoenix, the sample contains disproportionately more blacks and less educated young individuals. Using data from the CPS for year 2001, blacks and high school graduates represent 60% and 80% of the total youth population in Philadelphia (18 to 25 years old), respectively, against 70% and 39% in the sample. In Phoenix, Hispanics and high school graduates represent 66% and 73% of the total youth population, respectively, against 60% and 14% in the sample.

[9]Bushway and Sweeten (2007) document that ex-felons are barred from up to 800 different occupations across the United States.

[10]The overlap between legal and illegal work in the United States was also documented by Freeman (2000) using data from the NLSY.

Table 3.1: Pathways to Desistance - Descriptive Statistics

Variable	**Mean**
Black*	0.378
	(0.485)
Hispanic*	0.376
	(0.485)
White*	0.214
	(0.411)
Philadelphia*	0.471
	(0.500)
Age at Labour Market Entry*	18.899
	(1.494)
High School Degree*	0.245
	(0.430)
Income Crime Monthly Rate	0.106
	(0.308)
Legal Employment Monthly Rate	0.344
	(0.475)
Legal Experience	13.476
	(15.574)
Income Crime Experience	5.502
	(7.980)
Accumulated Criminal Records	5.217
	(3.553)
In Probation	0.171
	(0.376)
Number of Individuals	527

Notes:

* Indicates variables that do not vary over time. Summary statistics for these variables are calculated using only the survey at the time individuals stop attending school.

1. Standard deviations are reported below the mean in parenthesis.

2. Income crime and legal experience are measured in months and they ignore pre-survey experience. Accumulated criminal records are the sum of official arrests, including arrests that occurred before the survey.

might be particularly important for firms filling managerial positions or positions where monitoring is imperfect, further reducing the types of jobs available to individuals with evident criminal past.

Table 3.2: Pathways to Desistance - Distribution of Legal Workers by Type of Job

Type of Job	**Share**
Manual Occupations (e.g., Cutting Grass, Security Guard)	0.319
	(0.466)
Skilled Occupations (e.g., Carpentry)	0.304
	(0.460)
Restaurant Worker	0.154
	(0.361)
Office Work/Clerical/Telemarketing	0.061
	(0.238)
Retail/Cashier	0.071
	(0.257)
Managerial/Administrative	0.060
	(0.237)
Babysitting/Child Care	0.010
	(0.099)
Technical/Professional (e.g., Medical Assistant, Newspaper Reporter)	0.012
	(0.108)
Other	0.010
	(0.099)

Notes:
1. Standard deviations are reported below the mean in parenthesis.
2. The job type categories are those reported in the survey.

Table 3.3 displays the monthly average duration of non-employment (i.e., periods where the individual is not legally employed, participating in income crime, or incarcerated), legal jobs, and income crime in the sample. The short duration of legal jobs provides further evidence that disadvantaged young individuals are mostly employed in low-quality jobs. The average duration of legal jobs is 7.3 months. Table 3.3 also breaks down the average duration of legal jobs by hours worked and by type of job. Both full-time and part-time jobs have short average duration, with full-time jobs lasting only two months longer than part-time jobs, on average. There is, however, significant variation by type of job. Perhaps not surprisingly, managerial and administrative jobs display a longer average duration (14.6 months) than restaurant jobs (5.7 months) and manual occupations (6.3 months). Nevertheless, the duration of legal jobs in the sample is much shorter than what has been documented in the literature for all youths.[11]

Table 3.4 presents descriptive statistics on monthly legal earnings (i.e., one observation per

[11] Using data from the NLSY 1979, Bowlus, Kiefer, and Neumann (2001) document a mean job duration of 26.0 and 19.6 months for white and black young individuals, respectively, in their first jobs.

Table 3.3: Pathways to Desistance - Average Duration of Non-Employment, Legal Jobs, and Income Crime

Non-Employment	5.469
	(0.168)
Income Crime	4.108
	(0.096)
Legal Job	7.261
	(0.202)
Average Legal Job Duration By Hours Worked	
Part-Time	5.560
	(0.216)
Full-Time	7.707
	(0.151)
Average Legal Job Duration By Type of Legal Job	
Manual Occupations	6.351
	(0.191)
Skilled Occupations	9.006
	(0.222)
Restaurant Worker	5.720
	(0.083)
Office Work/Clerical/Telemarketing	6.207
	(0.310)
Retail/Cashier	7.180
	(0.213)
Managerial/Administrative	14.643
	(0.500)
Babysitting/Child Care	12.333
	(0.167)
Technical/Professional	11.286
	(0.286)

Notes:

1. Durations are in months and include censored spells.
2. Censoring rates are reported in parenthesis.
3. Non-employment refers to periods where the individual is not legally employed, participating in income crime or in prison.
4. The job type categories are those reported in the survey.

month with multiple observations for each legally employed individual). Average earnings in the legal labour sector amount to $1,438/month. This figure is substantially lower than the average monthly earnings of young individuals in their first job (Bowlus, Kiefer, and Neumann, 2001), but two times larger than the average monthly earnings of a full-time minimum-wage worker in Philadelphia and Phoenix in 2001.[12] There is, not surprisingly, significant variation in monthly earnings by type of job. Retail and restaurant jobs display the lowest monthly earnings, while clerical, managerial, and administrative jobs feature average salaries that are at least 33% higher than the average earnings in the lowest paying jobs.

Table 3.4: Pathways to Desistance - Average Monthly Legal Earnings

Monthly Legal Earnings	1,438.5
	(536.1)
Average Monthly Legal Earnings by Type of Job	
Manual Occupations	1,449.3
	(531.2)
Skilled Occupations	1,651.1
	(472.7)
Restaurant Worker	1,133.9
	(497)
Office Work/Clerical/Telemarketing	1,447.6
	(490.2)
Retail/Cashier	1,151.0
	(453.7)
Managerial/Administrative	1,472.4
	(488.8)
Babysitting/Child Care	1,799.7
	(720.3)
Technical/Professional	1,657.1
	(486.8)

Notes:
1. Standard deviations are reported below the mean in parenthesis.
2. Legal earnings are monthly and are expressed in 2000 US dollars.
3. The job type categories are those reported in the survey.

In Table 3.5, I further analyze earnings in the legal labour sector and explore the existence of returns to experience. I use the natural logarithm of reported hourly wages in the legal labour sector as the dependent variable. The wage regressions include individual fixed effects and controls for age, type of legal job, an indicator for part-time jobs, the accumulated criminal

[12] Bowlus, Kiefer, and Neumann, 2001 report average weekly earnings of $291 and $255 for whites and blacks, respectively, in the first jobs of young individuals. Using the year 2000 CPI to adjust the figures, average monthly earnings amount to $1,651 and $1,447 for whites and blacks, respectively. The 2001 state minimum wage was $5.15 both in Pennsylvania and Arizona, which corresponds to monthly earnings of $775 for a full-time minimum-wage job.

record, and an indicator for current participation in the criminal sector.[13] Legal labour market experience is measured in months and it ignores pre-survey experience. The results provide evidence of small returns to experience in the legal labour sector. An extra month of legal labour market experience increases hourly earnings by 0.6%. This is consistent with the fact that these are low-quality jobs, which I expect to feature small or no returns to experience. The results also confirm that there is significant dispersion in earnings across occupations, and that hourly wages in part-time jobs are 5.4% lower than in full-time jobs. Lastly, current participation in crime and accumulated criminal records play no significant role in earnings. These variables may still play an important role in the rate at which these individuals find jobs in the legal labour sector. This is explored in the next section.

The analysis indicates that youth previously involved in serious criminal activities face low-quality employment opportunities in the legal labour sector: legal jobs that do not last long, offer low wages, and are disproportionately in low-skilled occupations. I now turn to the alternative employment sector: the criminal sector. Table 3.6 displays the cross-section distribution of individuals engaged in income crime across types of crime. Drug-related crimes are the most prevalent among income crimes, followed by stealing merchandise and selling stolen property. In more than 70% of the individual-month observations in which an individual is participating in income crime, he is selling drugs. Individuals spend, on average, 4.1 months engaged in income crime (Table 3.4). This might seem short. However, the average duration of income crime is somewhat similar to the average duration of legal jobs suggesting that employment in the legal labour sector and the criminal sector is not that different. Furthermore, this indicates that these young individuals participate in the criminal sector for consecutive months.

Table 3.7 provides descriptive statistics on monthly criminal earnings (i.e., one observation per month with multiple observations for each individual engaged in income crime). There are several statistics that I want to stress. First, average earnings in the criminal sector amount to $3,654/month. There is also large variability in the distribution of criminal earnings relative to monthly legal earnings. Second, average monthly criminal earnings differ significantly depending on whether the individual engaged in income crime also has a job in the legal labour sector. Individuals who participate exclusively in income crime make roughly two times more than individuals who participate simultaneously in the criminal and legal labour sectors. One possible explanation is that individuals who have a legal job devote less hours to the criminal sector relative to individuals who concentrate exclusively on income crime.[14] Lastly, there is an evident earnings premium in the criminal sector, relative to the legal labour sector. Average monthly earnings in the criminal sector are almost three times higher than the mean reported legal earnings. Not surprisingly, some of the gap is driven by a large right tail, although a comparison of the median rates also reflects the earnings premium ($1,910/month versus $1,375/month). This characteristic makes the criminal sector quite attractive, and may partly explain why some individuals choose to participate in the criminal sector despite the risk of incarceration.

[13]The accumulated criminal record is the sum of official arrests, including arrests that occurred before the survey.

[14]The survey does not collect data on hours devoted to crime in a given day or week.

Table 3.5: Estimated Parameters from Legal Earnings Regressions

Variable	Dependent Variable Ln Hourly Wage (Average Hourly Wage = 9.44)
Age	-0.006
	(0.007)
Legal Labour Market Experience	0.006***
	(0.001)
Job Type: Retail/Cashier	0.084***
	(0.018)
Job Type: Babysitting/Child Care	0.430***
	(0.058)
Job Type: Skilled Occupations	0.250***
	(0.014)
Job Type: Manual Occupations	0.192***
	(0.013)
Job Type: Office Work/Clerical/Telemarketing	0.174***
	(0.019)
Job Type: Managerial/Administrative	0.255***
	(0.019)
Job Type: Technical/Professional	0.387***
	(0.036)
Currently Engaged in Income Crime	0.011
	(0.012)
Part-Time Job	-0.054***
	(0.009)
Accumulated Criminal Records	0.006
	(0.004)
Constant	1.942***
	(0.116)
Number of Observations	8,249
R-squared	0.148
Number of Individuals	399

Notes:

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

2. The excluded category for the type of job dummies is restaurant worker.

3. Experience is measured in months and it ignores pre-survey experience. Accumulated criminal records are the sum of official arrests, including arrests that occurred before the survey.

4. Hourly wages are expressed in 2000 US dollars.

Table 3.6: Pathways to Desistance - Distribution of Income Crime by Type of Crime

Type of Crime	**Share**
Sell Drugs	0.723
	(0.448)
Steal Merchandise	0.297
	(0.457)
Buy/Sell Stolen Goods	0.336
	(0.472)
Other	0.063
	(0.243)

Notes:
1. Standard deviations are reported below the mean in parenthesis.
2. An individual can be engaged in multiple income crime categories.
3. Other income crime includes gambling and prostitution.

Table 3.7: Pathways to Desistance - Average Monthly Criminal Earnings

Monthly Criminal Earnings	3,653.9
	(4585.4)
Monthly Criminal Earnings by Legal Employment Status	
No Legal Job	4,112.5
	(4767.2)
Legal Job	2,098.7
	(3488.5)

Notes:
1. Standard deviations are reported below the mean in parenthesis.
2. Criminal earnings are monthly and are expressed in 2000 US dollars.

In Table 3.8, I explore the existence of returns to experience in the criminal sector, as well as the role of gang participation and non-income crime on criminal earnings. I use the natural logarithm of reported weekly criminal earnings as the dependent variable in the regressions. The regressions include individual fixed effects and controls for age, number of income crime types the individual is involved in, an indicator for current participation in the legal labour sector, and the accumulated criminal record.[15] Income crime experience is measured in months and it ignores pre-survey experience. Indicators for gang and non-income crime involvement during the income crime spell are also included to evaluate their impact on earnings. The results in column (1) indicate that there are no returns to experience in income crime. In column (2), I take a closer look at the role of experience in the criminal sector by also accounting for experience in non-income crime (e.g., violent crime). While experience in income crime is not relevant for criminal earnings, the results provide suggestive evidence of returns to experience in non-income crime. An additional month of non-income crime experience increases weekly criminal earnings by 1.1%. Current participation in non-income crime and accumulated criminal records are not significantly related to the level of criminal earnings. Lastly, the results suggest that being part of a gang is negatively correlated with criminal earnings. One potential explanation is that gangs have a revenue-sharing scheme, in which total gains are shared among gang members (Chang, Lu, and Wang, 2013; Levitt and Venkatesh, 2000).

Overall, the results indicate that young individuals previously involved in serious criminal activities face two poor-quality employment alternatives in the criminal and legal labour sectors, which present small or no returns to experience. How individuals move across sectors presumably depends on how quickly they get opportunities in either sector. A closer look at the average duration of non-employment (i.e., the time an individual spends without employment in either sector) can provide suggestive evidence on this. The figures in Table 3.3 indicate that non-employment lasts, on average, 5.5 months. This average duration refers to periods when the individual is not working in either sector. Unemployment duration, as it is usually measured in the search literature (i.e., the time between legal employment spells), averages 6.6 months. These figures suggest that search frictions may be quite prevalent for disadvantaged young individuals. In particular, opportunities in the criminal and legal labour sectors may not be readily available, forcing them to wait until they face an attractive enough employment opportunity. In the next section, I further explore what are the key factors behind the transitions across sectors.

3.4 Transitions between the Criminal and Legal Labour Sector

In this section, I explore the transitions between non-employment, legal employment and income crime using the data described in Section 3.2. For this purpose, I start by explaining how the transitions between sectors are determined and then present results from a Mixed Propor-

[15] The accumulated criminal record is computed as the sum of official arrest records. It includes official arrests that occurred before the survey.

Table 3.8: Estimated Parameters from Criminal Earnings Regressions

Variable	**Dependent Variable Ln Weekly Wage (Average Weekly Wage = 1,025.5)**	
	(1)	**(2)**
Age	0.046	0.035
	(0.033)	(0.034)
Income Crime Experience	0.001	-0.005
	(0.005)	(0.005)
Currently Have a Legal Job	-0.067	-0.058
	(0.052)	(0.052)
Currently Engaged in Non-Income Crime	0.055	0.059
	(0.048)	(0.048)
Belongs to a Gang	-0.186*	-0.170*
	(0.106)	(0.106)
Number of Crime-Types is Engaged In	-0.018	-0.017
	(0.045)	(0.045)
Non-Income Crime Experience		0.011*
		(0.006)
Constant	5.343***	5.499***
	(0.619)	(0.624)
Number of Observations	1,433	1,433
R-squared	0.013	0.013
Number of Individuals	165	165

Notes:

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

2. Experience is measured in months and it ignores pre-survey experience. Accumulated criminal records are the sum of official arrests, including arrests that occurred before the survey.

3. Weekly wages are expressed in 2000 US dollars.

tional Hazards Competing Risks Model.

Monthly histories are constructed according to the following rules. Based on the major activity occurring during a particular month, an individual could be in one of the following five states: incarcerated, non-employed, employed in the legal labour sector, engaged in income crime, or both legally employed and engaged in income crime (employed/crime). The individual is incarcerated if he spends more than 15 days in jail, detention, or prison in that month.[16] Otherwise, the individual is non-employed, legally employed, engaged in income crime, or employed/crime. Non-employment is a residual state. An individual is classified as non-employed if he is not engaged in income crime, legally employed, or incarcerated in a given month. Note that this definition is different from the standard definition of unemployment, in which participation in the criminal sector is ignored. This distinction will be key in the analysis of transitions across sectors.

The monthly transitions are determined based on the five possible states (i.e., non-employment, legal employment, income crime, employment/crime, and incarceration). An individual makes a non-employment-to-legal employment transition if he is non-employed in the current month and legally employed in the next month. An individual makes a non-employment-to-crime transition if he is non-employed in the current month and engaged in income crime in the next month. Lastly, he makes a non-employment-to-incarceration transition if the individual is non-employed in the current month and incarcerated in the next month.[17] Monthly transitions for individuals legally employed, engaged in income crime, employed/crime, or incarcerated, are defined analogously.[18]

Table 3.9 illustrates the monthly transitions. There are particular transitions that I want to emphasize. First, the probability of engaging in income crime is slightly larger for non-employed individuals relative to individuals employed in the legal labour sector (3.8% against 3.1%). Second, for individuals currently engaged in income crime, the probability of a transition into legal employment is lower relative to non-employed individuals (4.5% against 8.2%). One potential explanation for these differences is that there is a negative effect of legal (criminal) earnings on the transition to income crime (legal employment). Nevertheless, these results do not control for any heterogeneity across individuals that could be driving this relationship (e.g., skills, education). I next attempt to understand the key factors behind these transitions.

[16]Only in 7.4% of the individual-month observations in which the individual reports to be incarcerated, the stay is shorter than 15 days.

[17]Non-employment-to-employment/crime and incarceration-to-employment/crime transitions represent less than 1% of all transitions corresponding to non-employment and incarceration spells. As a result, they are coded as missing transitions. Consistent with the data, these transitions are zero probability events in the model described in Chapter 4,

[18]One potential concern when using monthly transitions is that an individual may be categorized as employed/crime on a given month when he is really transitioning either from legal employment to crime, or viceversa. For example, individual A is participating in the criminal sector in month 1, participating in the legal and criminal sectors in month 2, and participating only in the legal labour sector in month 3. It is likely that the individual is not really participating simultaneously in the two sectors in month 2, but it is instead a consequence of data aggregation. To avoid this miss classification, the legal employment-to-crime/employment-to-crime transitions, where the middle state holds for exactly one month and the individual works for two weeks or less in the legal labour sector in that month, are re categorized as legal employment-to-crime transitions directly. Crime-to-employment/crime-

Table 3.9: Pathways to Desistance - Monthly Transitions

From Non-Employment to		From Incarceration to	
Non-Employment	0.845	Non-Employment	0.041
	(0.362)		(0.198)
Legal Employment	0.082	Legal Employment	0.010
	(0.274)		(0.102)
Income Crime	0.035	Income Crime	0.012
	(0.184)		(0.107)
Incarceration	0.038	Incarceration	0.937
	(0.192)		(0.243)
From Legal Employment to		**From Employment/Crime to**	
Non-Employment	0.061	Non-Employment	0.009
	(0.240)		(0.094)
Legal Employment	0.899	Legal Employment	0.180
	(0.301)		(0.385)
Income Crime	0.007	Income Crime	0.069
	(0.084)		(0.254)
Incarceration	0.009	Incarceration	0.023
	(0.092)		(0.149)
Employment/Crime	0.024	Employment/Crime	0.719
	(0.152)		(0.450)
From Income Crime to			
Non-Employment	0.114		
	(0.318)		
Legal Employment	0.016		
	(0.126)		
Income Crime	0.739		
	(0.440)		
Incarceration	0.102		
	(0.303)		
Employment/Crime	0.029		
	(0.168)		

Notes:

1. Standard deviations are reported below the mean in parenthesis.

2. Non-employment-to-employment/crime and incarceration-to-employment/crime transitions represent less than 1% of all transitions corresponding to non-employment and incarceration spells. As a result, they are coded as missing transitions.

3.4.1 Mixed Proportional Hazards Competing Risks Model

I analyze the transitions between sectors using the Mixed Proportional Hazards Competing Risks Model. Consider an individual in spell s for t months (e.g., a non-employment spell that lasts t months). All individual differences in the exit rate for destination j (e.g., legal employment) at time t can be characterized by observed characteristics x and unobserved characteristics v_j. This implies that the observed and unobserved characteristics of the individual shift the hazard rate for destination j in a proportional manner irrespective of the time elapsed since the start of the spell. For simplicity, I assume that there are two possible destinations for exit (e.g. $j = A, B$). The exit rates for destinations A and B conditional on x, v_A, and v_B are denoted by $\theta_A(t|x, v_A, v_B)$ and $\theta_B(t|x, v_A, v_B)$, respectively. These rates are assumed to have a Mixed Proportional Hazard specification,

$$\begin{aligned} \theta_A(t|x, v_A, v_B) &= \lambda_A(t)\phi_A(x)v_A \\ \theta_B(t|x, v_A, v_B) &= \lambda_B(t)\phi_B(x)v_B \end{aligned}$$

in which $\lambda_A(t)$ and $\lambda_B(t)$ represent the baseline hazards, v_A and v_B are unobserved correlated random variables that are distributed independently of x, and $\phi_j(x) = \exp(x'\beta_j)$ for $j = (A, B)$. Since I have access to multiple spells of type s for a given individual, I assume that the duration of each spell of type s is governed by the same probability laws and is affected by the same unobserved explanatory variables. This implies that the unobservables are fixed across spells of type s for a given individual. Furthermore, I assume that the intervening spells between any two spells of type s for a single individual are independent of the unobserved characteristics.

I assume an exponential specification for the baseline hazards and a bivariate distribution with unrestricted mass point locations for the joint cumulative distribution function of the unobservables. Furthermore, I assume that x does not vary over the spell. Abbring and Van den Berg (2003) show that all components of the model are identified. They also extend the identification analysis to the case with multiple-spell data and show that this type of data allows for identification under much weaker conditions than does single-spell type of data.

For estimation, I use the spell data summarized in Table 3.10. The data comprise 1,291 non-employment spells, 1,252 legal job spells, 474 income crime spells, 280 crime/employment spells, and 621 incarceration spells.[19] Tables 3.11-3.13 show the results of the model for non-employment, income crime, and legal employment spells. Each column explores how different variables affect the hazard rate of leaving the current state into a particular state (e.g., from income crime to legal employment). Depending on the current state, there are different transitions the individual can make. For example, non-employed individuals can transition into legal employment, income crime, or incarceration. Instead, individuals who are currently engaged

to-legal employment transitions are recategorized analogously.

[19]Note that non-employment-to-employment/crime and incarceration-to-employment/crime transitions represent less than 1% of all transitions corresponding to non-employment and incarceration spells. As a result, they are coded as missing transitions. Also, legal employment-to-legal employment transitions and employment/crime-to-employment/crime transitions entail changes in the employer. Lastly, income crime-to-income crime transitions are zero probability events by construction.

Table 3.10: Pathways to Desistance - Spell Data Sample Means

Average Duration		**Average Duration**	
Non-Employment	5.469	Incarceration	11.962
	(6.509)		(13.997)
Legal Employment	6.306	Employment/Crime	3.060
	(7.485)		(3.305)
Income Crime	3.614		
Flow from Non-Employment to		**Flow from Incarceration to**	
Legal Employment	0.528	Non-Employment	0.650
	(0.499)		(0.478)
Income Crime	0.226	Legal Employment	0.166
	(0.418)		(0.372)
Incarceration	0.246	Income Crime	0.184
	(0.431)		(0.388)
Flow from Legal Employment to		**Flow from Employment/Crime to**	
Non-Employment	0.432	Non-Employment	0.029
	(0.496)		(0.167)
Legal Employment	0.292	Legal Employment	0.586
	(0.455)		(0.494)
Income Crime	0.050	Income Crime	0.225
	(0.218)		(0.419)
Incarceration	0.060	Incarceration	0.074
	(0.238)		(0.262)
Employment/Crime	0.166	Employment/Crime	0.086
	(0.373)		(0.281)
Flow from Income Crime to			
Non-Employment	0.437		
	(0.497)		
Legal Employment	0.061		
	(0.24)		
Incarceration	0.391		
	(0.488)		
Employment/Crime	0.111		
	(0.314)		

Notes:

1. Standard deviations are reported below the mean in parenthesis.

2. Durations are in months and include censored spells.

3. The transition probabilities sum to one since I only consider completed (i.e., uncensored) spells for the calculation.

4. Non-employment-to-employment/crime and incarceration-to-employment/crime transitions represent less than 1% of all transitions corresponding to non-employment and incarceration spells. As a result, they are coded as missing transitions. Legal employment-to-legal employment transitions and employment/crime-to-employment/crime transitions entail changes in the employer. Income crime-to-income crime transitions are zero probability events by construction.

in income crime can transition into non-employment, legal employment, employment/crime, or incarceration. The possible transitions for each state are specified in the tables.[20] The coefficients represent the effect on the log hazard rate. In an attempt to account for individual heterogeneity, each regression includes controls for age, race, location, an indicator for high school graduation, legal labour sector experience, and experience in income crime.[21] I also include as regressors the estimated cognitive and social/emotional skill factors from Chapter 2. Since probation and accumulated criminal records are expected to affect the choice of participating in income crime and/or taking a legal job, I also add these variables as regressors. Lastly, I include an indicator for participation in non-income crime to evaluate if it has any effect on the transitions after accounting for engagement in income crime.[22] As I do not normalize the scale of the unobservables, I exclude a constant from the regressors and one category from each set of dummies.

I start by analyzing the main drivers behind the transitions into income crime. To this end, I concentrate on non-employment and legal employment spells. The results in Tables 3.11 and 3.12 indicate that, while cognitive skills are not significantly related to income crime participation, social/emotional skills show a strong negative correlation with it. A one standard deviation increase in social/emotional skills leads to a decrease in the hazard rate of income crime of 41.6% and 49.2% for non-employed and legally employed individuals, respectively. These results are consistent with the findings in Chapter 2 for violent, drug-related, and property crime.

I also account for the role of experience in income crime. The results suggest that an additional month of experience in the criminal sector increases the hazard rate of participating in income crime 6% to 7%. Whereas the results in Section 3.3 provided no evidence of returns to experience in income crime, criminal experience significantly affects transitions. This suggests that experience in income crime might be related to participation in income crime in other dimensions. For instance, more criminal experience may increase the frequency of opportunities in the criminal sector.

Since probation is expected to directly alter the hazard rate of entering income crime, for example via more more supervision by a probation officer, I also include an indicator variable for probation status. The results suggest that being in probation reduces the hazard rate of income crime by 46.4% for non-employed individuals. The effect is also negative for legally employed individuals, although it is not precisely estimated.

The descriptive analysis of transitions in Table 3.9 suggested that earnings in the legal

[20]In order to focus on the transitions from legal employment to income crime, the legal employment-to-income crime and legal employment-to-employment/crime transitions are pulled together for the analysis on legal employment spells (see Table 3.12). Likewise, in order to focus on the transitions from income crime to legal employment, the income crime-to-legal employment and income crime-to-employment/crime transitions are pulled together for the analysis on income crime spells (see Table 3.13).

[21]For time-varying variables, like age and experience, I take the average over the spell. Experience in income crime and legal employment are measured in months and they ignore pre-survey experience.

[22]Accumulated criminal records are the sum of official arrests, including arrests that occurred before the survey. The role of probation and non-income crime is accounted for by including indicators for whether the individual participated in non-income crime and was in probation at some point during the spell.

Table 3.11: Estimated Parameters from Mixed Proportional Hazards Competing Risks Model - Non-Employment Spells

Variable	(1) Legal Employment	(2) Income Crime	(3) Incarceration
Age	-0.167***	-0.086	0.055
	(0.048)	(0.075)	(0.049)
Black	0.020	-0.463	0.832**
	(0.206)	(0.317)	(0.359)
Hispanic	-0.001	-0.500**	0.496*
	(0.159)	(0.248)	(0.301)
Philadelphia	-0.416***	-0.082	-0.649***
	(0.162)	(0.266)	(0.261)
Cognitive Skills	0.396	0.419	0.320
	(0.270)	(0.404)	(0.391)
Social/Emotional Skills	0.314**	-0.945***	-0.181
	(0.160)	(0.268)	(0.199)
High School Graduate	-0.002	-0.065	0.133
	(0.164)	(0.261)	(0.244)
Legal Labour Market Experience	0.026***	-0.031***	-0.005
	(0.007)	(0.013)	(0.010)
Income Crime Experience	-0.002	0.060***	0.024*
	(0.008)	(0.012)	(0.013)
Accumulated Criminal Records	-0.032	0.036	0.152***
	(0.021)	(0.030)	(0.030)
In Probation	0.054	-0.464**	-0.085
	(0.133)	(0.200)	(0.196)
Engaged in Non-Income Crime	-0.708***	-0.289	-0.188
	(0.143)	(0.257)	(0.180)
Number of Observations	1,055	1,055	1,055
Number of Failures	467	193	217

Notes:

1. The coefficients represent the effect on the log hazard rate. Columns (1), (2), and (3) present the results corresponding to the hazard rate of legal employment, income crime, and incarceration, respectively, for non-employed individuals.

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

3. Age and experience are averages over spells. Experience is measured in months and it ignores pre-survey experience. Accumulated criminal records are the sum of official arrests, including arrests that ocurred before the survey. The cognitive skills and non-cognitive (social/emotional) skills estimates come from the factor analysis in Chapter 2.

Table 3.12: Estimated Parameters from Mixed Proportional Hazards Competing Risks Model - Legal Employment Spells

Variable	(1) Non-Employment	(2) Legal Employment	(3) Income Crime	(4) Incarceration
Age	-0.019	-0.139***	-0.138***	0.191***
	(0.019)	(0.027)	(0.040)	(0.071)
Black	-0.167	-0.296	-0.143	0.908
	(0.207)	(0.304)	(0.378)	(0.662)
Hispanic	-0.137	-0.205	-0.173	0.537
	(0.162)	(0.239)	(0.301)	(0.510)
Philadelphia	0.437***	-0.305	0.289	-0.638
	(0.170)	(0.252)	(0.321)	(0.659)
Cognitive Skills	-0.110	0.006	0.621	0.524
	(0.332)	(0.401)	(0.487)	(0.84)
Social/Emotional Skills	-0.081	-0.084	-1.117***	0.327
	(0.188)	(0.254)	(0.387)	(0.527)
High School Graduate	-0.350**	-0.290	-0.040	-0.085
	(0.167)	(0.220)	(0.262)	(0.643)
Legal Labour Experience	-0.020***	0.014***	0.001	-0.034**
	(0.007)	(0.001)	(0.001)	(0.017)
Income Crime Experience	-0.011	-0.002	0.069***	-0.020
	(0.016)	(0.018)	(0.003)	(0.050)
ln(Monthly Legal Earnings)	-0.728***	-0.661***	-0.241***	-0.549
	(0.088)	(0.123)	(0.094)	(0.375)
Accumulated Criminal Records	0.004	0.019	-0.009	0.124*
	(0.027)	(0.030)	(0.045)	(0.071)
Engaged in Non-Income Crime	-0.508**	-1.038***	-0.310	0.219
	(0.226)	(0.267)	(0.253)	(0.598)
In Probation	-0.245	0.202	-0.118	1.462***
	(0.185)	(0.200)	(0.216)	(0.394)
Number of Observations	1,069	1,069	1,069	1,069
Number of Failures	405	276	204	57

Notes:

1. The coefficients represent the effect on the log hazard rate. Columns (1), (2), (3), and (4) present the results corresponding to the hazard rate of non-employment, legal employment (exclusive participation in the legal labour sector), crime (with or without employment in the legal labour sector), and incarceration, respectively, for individuals employed in the legal labour sector exclusively.
2. Standard errors are reported below the point estimates in parentheses. *** stands for p-value<0.01, ** stands for p-value<0.05, * stands for p-value<0.1.
3. Age, experience, and earnings are averages over spells. Experience is measured in months and it ignores pre-survey experience. Accumulated criminal records are the sum of official arrests, including arrests that ocurred before the survey. The cognitive skills and non-cognitive (social/emotional) skills estimates come from the factor analysis in Chapter 2.

labour sector may play an important role in the decision to participate in income crime. The results in Table 3.12 are consistent with this, as the sign of legal earnings is negative and significant for the hazard rate of entering income crime from legal employment. In particular, a 10% increase in monthly legal earnings reduces the hazard rate of entering income crime by 2.4%.

Lastly, since the analysis focuses on income crime, I also included an indicator for participation in non-income crime (e.g., violent crime) in the regressions for non-employment and legal employment spells, to evaluate if this activity plays any role in the transitions of individuals not currently engaged in income crime. The results indicate that participation in non-income crime does not affect the transition to income crime. Of particular interest are the transitions from non-employment and legal employment into incarceration. One could expect, for example, that non-employed individuals or legally employed individuals who currently participate in non-income crime are more prone to be incarcerated than those who do not engage in these type of crimes. The results suggest that, after accounting for participation in income crime, engagement in non-income crime does not play a significant role in the transitions to incarceration.

I now analyze the main factors behind the transition into the legal labour sector. I focus on non-employment and income crime spells (Tables 3.11 and 3.13). Consistent with what one would expect, I find that higher cognitive and social/emotional skills increase the hazard rate of taking a legal job. The results for non-employed individuals imply that a one standard deviation increase in social/emotional skills leads to an 13.8% increase in the hazard rate of taking a job in the legal labour sector. The effects of cognitive skills are imprecisely estimated.

I also examine the effect of legal labour market experience on the hazard rate of legal employment. Despite having a small effect on legal earnings, experience in the legal labour sector is strongly related to the transition to legal employment. I find that an additional month of legal labour market experience increases the hazard rate of legal employment by 2.6% for non-employed individuals. The effect is positive and not precise for the income crime-to-legal employment transition.

A somewhat surprising result is that the accumulated criminal records play no significant role on the transitions to the legal labour sector, since criminal records are viewed as a barrier to legal employment in the literature. One explanation for this finding is that the distribution of accumulated criminal record is shifted to the right, relative to the general youth population (i.e., everybody has been arrested at least once). This suggests that conditional on being arrested once, an additional arrest has no significant effect on legal employment.

With regards to the effect of criminal earnings on the hazard rate of legal employment, the results in Table 3.13 indicate that the probability that an income crime spell ends in a legal job significantly decreases with monthly criminal earnings. Specifically, a 10% increase in monthly criminal earnings reduces the hazard rate of an income crime-to-legal employment transition by 5.3%.

Altogether, there are some important messages from the analysis on transitions across sectors. First, I find that individual heterogeneity is strongly related to criminal and legal employ-

Table 3.13: Estimated Parameters from Mixed Proportional Hazards Competing Risks Model - Income Crime Spells

Variable	**(1) Non-Employment**	**(2) Legal Employment**	**(3) Incarceration**
Age	-0.013	-0.173	0.060
	(0.113)	(0.216)	(0.080)
Black	-0.821***	0.355	-0.309
	(0.347)	(1.355)	(0.361)
Hispanic	-0.004	-0.322	-0.298
	(0.389)	(0.667)	(0.322)
Philadelphia	0.439	-0.590	-0.193
	(0.380)	(1.430)	(0.338)
Cognitive Skills	0.878*	0.945	0.280
	(0.500)	(1.044)	(0.475)
Social/Emotional Skills	-0.186	-0.483	-0.026
	(0.415)	(0.704)	(0.265)
High School Graduate	-0.623	-0.162	0.428
	(0.418)	(0.580)	(0.305)
Legal Labour Market Experience	-0.036*	0.039	-0.004
	(0.020)	(0.029)	(0.013)
Income Crime Experience	0.001	-0.015	0.003
	(0.017)	(0.029)	(0.010)
ln(Monthly Criminal Earnings)	-0.377***	-0.531***	0.201***
	(0.109)	(0.171)	(0.086)
Accumulated Criminal Records	-0.087*	-0.036	0.063**
	(0.052)	(0.091)	(0.030)
Engaged in Non-Income Crime	-0.883***	-1.040*	-0.289
	(0.260)	(0.543)	(0.221)
In Probation	-0.263	0.358	0.163
	(0.282)	(0.531)	(0.225)
Belongs to a Gang	-0.212	0.110	0.368
	(0.413)	(2.102)	(0.307)
Number of Observations	238	238	238
Number of Failures	71	32	113

Notes:

1. The coefficients represent the effect on the log hazard rate. Columns (1), (2), and (3) present the results corresponding to the hazard rate of non-employment, legal employment (with or without participation in the criminal sector), and incarceration, respectively, for individuals participating in income crime exclusively.

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

3. Age, experience, and earnings are averages over spells. Experience is measured in months and it ignores pre-survey experience. Accumulated criminal records are the sum of official arrests, including arrests that ocurred before the survey. The cognitive skills and non-cognitive (social/emotional) skills estimates come from the factor analysis in Chapter 2.

ment behaviour. In particular, social/emotional skills are important drivers of the hazard rate of entering income crime, and both cognitive and social/emotional skills are key aspects of the hazard rate of entering legal employment. Second, while the analysis in Section 3.3 provided evidence of small returns to experience in the criminal and legal labour sectors, higher experience in a given sector plays a significant role for transitions to the corresponding sector. This indicates that experience is potentially less important for these two low-quality activities relative to high-skilled activities. Nevertheless, experience can still be relevant for other dimensions beyond earnings. Lastly, the criminal and legal labour sectors are strongly related. Overall, higher earnings in a given sector reduce the chances of accepting offers in the alternative employment sector. This suggests that income crime choice should not be analyzed in isolation from the legal employment, and viceversa. Since the criminal sector appears as an alternative to the legal labour sector, especially for disadvantaged young individuals, an appropriate approach is to study these two sectors jointly.

3.5 Conclusions

Youth unemployment and youth crime are two phenomenon that are disproportionately concentrated among disadvantaged young individuals. In this chapter, I characterize the criminal and legal labour sectors and empirically investigate the factors driving the transitions between sectors for a particular disadvantaged population group: young offenders. For this purpose, I use a unique dataset of young offenders: the Pathways to Desistance. The extensive data available on legal employment and crime participation allow me to provide a thorough description of what the labour market looks like for these individuals.

I find that young individuals previously involved in serious criminal activities face two poor-quality employment alternatives in the criminal and legal labour sectors. Legal jobs feature low wages, are in disproportionately low-skilled occupations, and do not tend to last. Income crime activity shows similar characteristics, although this sector offers an earnings premium to compensate for its risk. Consistent with their low quality, both sectors present small or no returns to experience.

The transitions across the criminal and legal labour sectors are strongly related to individual heterogeneity. In particular, I find an important role of social/emotional skills for income crime choice, and that both cognitive and social/emotional skills are key aspects of the legal employment choice. I show that, despite not being related to earnings in the corresponding sector, higher experience in a given sector plays a significant role in transitions. Finally, I document that the criminal and legal labour sectors are strongly related. For instance, higher criminal earnings reduce the chances of accepting legal job offers. A similar relation is found between current legal earnings and the hazard of entering income crime. Overall, this suggests that legal employment and crime choices should be studied jointly. Among disadvantaged young individuals, the criminal sector appears as an alternative to the legal labour sector. Therefore, legal employment or crime choices should not be analyzed without considering the alternative employment sector.

Lastly, it is important to stress that I study youth who have already committed somewhat serious criminal offenses. Although this is a particularly relevant group to study, as they represent a large proportion of youth crime and disproportionately contribute to youth unemployment, there are some caveats related to focusing on this specific group. One implication is that my results do not necessarily generalize to the youth population at large. In this sense, some of the factors that cause these individuals to transition into crime or legal employment may not be relevant for disadvantaged young individuals who have never participated in crime.

Bibliography

Abbring, Jaap H. and Gerard J. Van den Berg. 2003. "The Identifiability of the Mixed Proportional Hazards Competing Risks Model." *Journal of the Royal Statistical Society, Series B* 65 (3):701–710.

Agan, Amanda and Sonja Starr. 2017. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *Quarterly Journal of Economics* , forthcoming.

Bowlus, Audra J., Nicholas M. Kiefer, and George R. Neumann. 2001. "Equilibrium Search Models and the Transition from School to Work." *International Economic Review* 42 (2):317–343.

Bushway, Shawn D. 2004. "Labor Market Effects of Permitting Employer Access to Criminal History Records." *Journal of Contemporary Criminal Justice* 20 (3):276–291.

Bushway, Shawn D. and Gary Sweeten. 2007. "Abolish Lifetime Bans for Ex-Felons." *Criminology & Public Policy* 6 (4):697–706.

Cameron, Stephen V. and James J. Heckman. 1993. "The Nonequivalence of High School Equivalents." *Journal of Labor Economics* 11 (1, Part 1):1–47.

Chang, Juin-Jen, Huei-Chung Lu, and Ping Wang. 2013. "Search for a Theory of Organized Crimes." *European Economic Review* 62:130–153.

Eckstein, Zvi and Kenneth I. Wolpin. 1995. "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model." *The Review of Economic Studies* 62 (2):263–286.

Freeman, Richard B. 1991. "Crime and the Employment of Disadvantaged Youths." Working Paper 3875, National Bureau of Economic Research.

Freeman, Richard B. 2000. "Disadvantaged Young Men and Crime." In *Youth Employment and Joblessness in Advanced Countries*. University of Chicago Press, 215–246.

Freeman, Richard B. and David A. Wise. 1982. "The Youth Labor Market Problem: Its Nature Causes and Consequences." In *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*. University of Chicago Press, 1–16.

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

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

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

Levitt, Steven D. and Sudhir Alladi Venkatesh. 2000. "An Economic Analysis of a Drug-Selling Gang's Finances." *The Quarterly Journal of Economics* 115 (3):755–789.

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

Machin, Stephen and Costas Meghir. 2004. "Crime and Economic Incentives." *Journal of Human Resources* 39 (4):958–979.

Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108 (5):937–975.

Raphael, Steven. 2011. "Improving Employment Prospects for Former Prison Inmates: Challenges and Policy." In *Controlling Crime: Strategies and Tradeoffs.*, edited by Phillip J. Cook, Jens Ludwig, and McCrary Justin, chap. II. University of Chicago Press, 521–572.

Raphael, Steven and Melissa Sills. 2007. "Urban Crime, Race, and the Criminal Justice System in the United States." *A Companion to Urban Economics* :515–535.

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

Schnepel, Kevin T. 2016. "Good Jobs and Recidivism." *The Economic Journal* , forthcoming.

Uggen, Christopher, Mike Vuolo, Sarah Lageson, Ebony Ruhland, and Hilary K. Whitmam. 2014. "The Edge of Stigma: An Experimental Audit of the Effects of Low-Level Criminal Records on Employment." *Criminology* 52 (4):627–654.

Wolpin, Kenneth I. 1987. "Estimating a Structural Search Model: The Transition from School to Work." *Econometrica* 55 (4):801–817.

Chapter 4

A Search Model of Early Employment Careers and Youth Crime

4.1 Introduction

The school to work transition and the early career path can be a difficult one for young people, especially among the low-educated. These disadvantaged individuals tend to have low employment rates, experience long periods of unemployment before their first job, and face lower wages than their older counterparts (Wolpin, 1987; Eckstein and Wolpin, 1995; Bowlus, Kiefer, and Neumann, 2001). These circumstances motivate the public sector to assign many resources to training and education programs (e.g., the Job Corps), wage subsidies, and job search assistance programs (Katz, 1996; Heckman, LaLonde, and Smith, 1999), with the objective of improving the labour market outcomes of disadvantaged youth.

One reason why the public sector focuses on reducing unemployment and enhancing wages among disadvantaged young individuals is to prevent criminal activity.[1] The standard economic model of criminal behaviour predicts that an increase in job availability in the legal labour sector reduces the amount of time devoted to crime (Becker, 1968; Ehrlich, 1973). Moreover, the empirical evidence finds a strong association between unemployment and criminal activity (Raphael and Winter-Ebmer, 2001; Gould, Weinberg, and Mustard, 2002; Lochner, 2004), as well as low wages and crime (Grogger, 1998; Machin and Meghir, 2004; Kling, 2006).

Not surprisingly, a large fraction of the disadvantaged youth are engaged in crime. The literature has documented that crime is widespread among young low-educated males in poor urban areas (Freeman, 2000; Raphael and Sills, 2007). Many make money from illegal activities, and some even have a career in crime, spending most of their time engaged in criminal activities. Around 30% of low-educated young males in the United States reported an income from crime in 1980 (Lochner, 2004). The extensive involvement of young men in crime im-

[1]Another concern in the literature is that unemployment during youth may contribute to unemployment and low wage rates later in life (Freeman and Wise, 1982).

poses a cost on the victims of crime and it urges the government to devote significant resources to crime prevention. In 2003, the United States expenditures on police protection, judicial activities, and corrections amounted to a record \$185 billion, accounting for 7.2% of all State and local public expenditures (Hughes, 2006).

The linkages between the criminal and legal labour sectors were also documented in the Chapter 3. In particular, I showed that the probability of engaging in income crime is slightly larger for non-employed individuals relative to individuals employed in the legal labour sector. Second, for individuals currently engaged in income crime, the probability of a transition into legal employment is lower relative to non-employed individuals. Furthermore, higher criminal earnings reduce the chances of accepting legal job offers. A similar relation is found between current legal earnings and the hazard of entering income crime.

Despite the strong links between criminal activity and the legal labour sector, the youth labour market for disadvantaged individuals has been studied focusing primarily on the legal labour sector, ignoring the criminal sector. Neglecting this alternative employment sector may lead to inaccurate predictions of youth behaviour, especially among disadvantaged groups. Since crime provides an alternative to the legal labour sector, choices in the criminal sector may have an impact on choices in the legal labour sector, and the other way around. Young individuals that are making money in the criminal sector, for example, may not be searching for legal jobs.

Failing to account for the trade-offs between the criminal and legal labour sectors may also have implications for policy makers interested in boosting legal employment among disadvantaged youth. For instance, individuals who have a career in the criminal sector may respond to a policy that raises earnings in the legal labour sector (e.g., a wage subsidy) in a different manner than individuals who do not participate in crime, because their outside option is different. Active criminals presumably have less incentives to take a legal job, and as a consequence, what helps to boost youth legal employment at large, may not be as useful for increasing legal employment for disadvantaged youth who participate in crime.

In this chapter, I characterize both the legal labour sector and the criminal sector for disadvantaged youth in the context of a two-sector search model.[2] The search framework is well-suited to this type of study because it brings together some key features, like the long periods of non-employment and the large fraction of youth making money from crime, and ties them to frictions and earnings differences in the criminal and legal labour sectors. I model the legal labour sector using a standard setup with on-the-job search (Mortensen, 1986) in order to capture wage growth across legal job spells. I model participation in the criminal sector as self-employment since it is uncommon to have an employer in this sector, and I cannot easily establish a change in careers within the criminal sector as I can in the legal labour sector. The model includes criminal earnings shocks to capture the earnings dispersion within a crime spell. The criminal earnings shocks result in a new stream of criminal earnings and the option

[2]Two-sector search models have been used in the literature to study formal and informal labour sectors (Albrecht, Navarro, and Vroman, 2009; Cano-Urbina, 2015), wage-employment and self-employment (Lain, 2016; Narita, 2012), public and private labour sectors (Albrecht, Robayo-Abril, and Vroman, 2017), and dual labour markets (Albrecht and Vroman, 1992).

to exit crime. Consistent with what I documented empirically in Chapter 3, modeling crime in this way yields spells of criminal activity rather than isolated acts of crime. Lastly, since disadvantaged youth often participate in the criminal and legal labour sectors at the same time (Freeman, 2000), I extend the two-sector search model to allow individuals to simultaneously hold a legal job and participate in crime. While previous search models allow for multiple activities at the household level, none, to the best of my knowledge, allow for multiple activities at the individual level (i.e., multiple job holding).[3]

The framework closest to mine is that of Burdett, Lagos, and Wright (2003, 2004).[4] In their setup, unemployed and legally employed individuals receive offers from the legal labour sector and also encounter crime opportunities. A crime opportunity is a sporadic event that can end in jail. In their model, the legal labour sector has implications for the choice of whether to engage in crime, as individuals with higher wages are less likely to commit crimes due to higher opportunity costs of getting caught and sent to prison. However, there are no direct effects of the criminal sector on the legal labour sector, since crime is modeled as an instantaneous activity.[5] In comparison, my model allows for a much richer set of interactions between the two sectors, which are key to understanding how disadvantaged young individuals make choices across sectors. The criminal sector can affect the legal labour sector in two main ways. First, crime takes time that has to be re-allocated from leisure or legal employment. For large enough criminal earnings, individuals will quit their legal jobs to focus on the criminal sector. Second, crime is allowed to affect offer rates, destruction rates, and incarceration rates, to capture, for example, how criminal activity affects the possibilities of finding or keeping a legal job. In fact, I find that the arrival rate of legal job offers for individuals participating in crime is roughly half that of individuals who do not participate in crime.

Estimation of my model requires comprehensive information on participation in the criminal and legal labour sectors that enables the construction of legal job and crime spells. Therefore, the data on criminal and legal labour sectors participation needs to be collected as an event history, allowing me to determine whether an individual was engaged in a particular activity for a specific period of time. The model further requires detailed information on both criminal and legal earnings.

To meet these needs, I take advantage of a unique panel dataset, the Pathways to Desistance Study (PDS). The PDS is a multi-site longitudinal study of young offenders as they transition from adolescence into early adulthood in Philadelphia, Pennsylvania and Phoenix, Arizona. It was designed specifically to study questions related to the evolution of criminal behaviour,

[3]See for example, Dey and Flinn (2008); Guler, Guvenen, and Violante (2012), and Flabbi and Mabli (2018). These papers develop a household search model where none, one, or both spouses can be employed.

[4]Other theoretical models of crime and legal employment that use a search framework include Engelhardt, Rocheteau, and Rupert (2008), Huang, Laing, and Wang (2004), and Chang, Lu, and Wang (2013). Building on Burdett, Lagos, and Wright (2004, 2003), Engelhardt (2010) provides empirical estimates of the effects of employment frictions on crime, while Braun (2018) adapts the search model to quantify the effects of changing the minimum wage. Other models of crime and legal employment that do not use a search framework include İmrohoroğlu, Merlo, and Rupert (2004); Imai and Krishna (2004); Lochner (2004); Mocan, Billups, and Overland (2005), and Sickles and Williams (2008). For papers focusing specifically on youth see Munyo (2015) and Merlo and Wolpin (2015).

[5]There is still an opportunity cost due to the possibility of getting caught and sent to jail.

taking special care to also measure employment in the legal labour sector. The survey covers youth who were found guilty of a serious criminal offense committed between the ages of 14 and 18. These young individuals are a very disadvantaged group, with a large share of minorities and low-educated individuals. Each participant was followed for a period of seven years, which results in a comprehensive picture of life changes in a wide array of areas over the course of this time. The dataset contains a rich panel of information about decisions to participate in the legal labour sector and the criminal sector, as well as detailed information of transitions between the two sectors. As a result, the PDS is especially well-suited for estimating my model and helping to understand the interactions between the criminal and legal labour sectors.

Most studies that analyze the criminal and the legal labour sector jointly use more common datasets such as the National Longitudinal Survey of Youth (NLSY), which samples the population at large.[6] Instead, the PDS concentrates on young offenders. One advantage of using this survey is that I can focus on a group that is rarely studied but contributes significantly to aggregate crime rates. Another advantage is that I can concentrate on a group of individuals who are fairly active in both the criminal and legal labour sectors, which helps to develop a better understanding of the interactions across the two sectors. The PDS also has very detailed information regarding participation in the criminal sector that allows me to construct complete spells in the criminal sector, such as earnings data, and the number of weeks participating in crime in a given month.[7]

I estimate my model separately for each location using Indirect Inference. As a preview of my results, I find that individuals in the PDS face considerable search frictions in the legal labour sector. I estimate that individuals in Philadelphia who do not have a legal job and do not participate in crime, receive a legal job offer every 14 months, on average, which is twice as long as is usually estimated in the literature for youth.[8] These individuals are offered low-quality legal jobs that are characterized by short durations, low earnings, and large destruction rates. The criminal sector offers an earnings premium relative to the legal labour sector, which makes it an attractive alternative. Nevertheless, crime brings a higher probability of incarceration and fewer opportunities in the legal labour sector. I find that there are sizable interactions across sectors, and that policies in one sector can have important effects on the other sector. I provide evidence that policies targeting the legal labour sector (e.g., wage subsidy) can reduce crime and boost legal employment among disadvantaged youth. Furthermore, a policy that reduces the arrival rate of crime opportunities (e.g., via increasing the number of police), com-

[6] Some papers that use the NLSY79 to study the criminal and legal labour sectors include Lochner (2004); Engelhardt (2010); Grogger (1998). Merlo and Wolpin (2015) use the NLSY97. Another dataset often used in the literature is the Philadelphia Birth Cohort Study (Imai and Krishna, 2004; Sickles and Williams, 2008).

[7] Both the NLSY79 and NLSY97 record detailed data on participation in the legal labour sector. Regarding crime data, the NLSY79 collects, only at one specific survey, a number of questions about participation in crime and delinquent activities. While the NLSY97 collects information on crime participation at every survey, the data do not permit the creation of an event history at a monthly level since the survey gathers information on the number of times individuals participated in a given criminal activity since the date of last interview (i.e., during the last year). Being able to construct a monthly event history (or smaller frequency) is necessary to fully capture the transitions and interactions across sectors.

[8] For empirical estimates of standard search models of the youth legal labour market see the survey by Eckstein and Van den Berg (2007).

pared to extending the average sentence length, has the advantage of reducing crime without generating large increases in the incarcerated population.

The rest of the chapter is organized as follows. In Section 4.2, I develop the model. In Section 4.3, I present the estimation results from the model. In Section 4.4, I discuss some policy simulations. Finally, Section 4.5 concludes.

4.2 Model

The framework is a two-sector search model. The first sector is a standard legal labour sector (Mortensen, 1986). The second sector is a criminal sector modeled as self-employment. This is consistent with the nature of crime, where individuals usually work for themselves rather than for an employer. Following much of the empirical search literature, I adopt a partial equilibrium framework (Mortensen and Pissarides, 1999; Mortensen, 1986).

The economy is populated by a continuum of homogeneous, risk-neutral, and infinitely-lived workers, who maximize the discounted stream of expected lifetime utility. Time is continuous and individuals discount the future with interest rate r. The state variables upon which workers make decisions include the employment state, legal earnings w, and criminal earnings y. At each point in time individuals can be non-employed, legally employed, devoted to income crime, participating in both sectors, or incarcerated. Let the value functions of each state be represented by V^{ne}, $V^{e}(w)$, $V^{c}(y)$, $V^{ec}(w, y)$, and J, respectively.

I extend the two-sector search model to allow for simultaneous participation in both sectors. To this end, I introduce an intensive margin of labour supply. Hours worked in the criminal and legal labour sectors are denoted as h_c and h_e, respectively, and $l \in (0, 1)$ stands for leisure, where $l = 1 - h_e - h_c$. Hours worked in a given sector are equal to $\frac{2}{3}$ and $\frac{1}{3}$ when the activity is full-time and part-time, respectively. I assume that legal employment and income crime are full-time activities when the individual is fully devoted to either sector. Therefore, the individual partly benefits from leisure if he only participates in one sector ($l = \frac{1}{3}$). If the individual participates in both sectors simultaneously, the legal employment and income crime are full-time and part-time activities, respectively.[9] In this case, the individual does not benefit from leisure ($l = 0$).

Individuals draw from a legal earnings offer distribution $F(w)$ with mean μ_w and variance σ_w^2. Monthly legal earnings equal the product of w times the hours worked in the legal labour sector. Legal job offers arrive at rate λ. If an individual accepts the offer, the monthly legal earnings remain constant for the duration of the job. As is standard in the literature, the model allows for on-the-job search in the legal labour sector to account for earnings growth across legal job spells.

[9]This is consistent with the facts observed in the data used for the empirical analysis: 73% of the individual-month observations in which the individual is employed in the legal labour sector, he does so on a full time basis. The survey does not collect data on hours devoted to income crime.

In addition, individuals take draws from a criminal earnings offer distribution $M(y)$ with mean μ_y and variance σ_y^2. Monthly criminal earnings equal the product of y times the hours devoted to income crime. The opportunities in the criminal sector arrive at rate η. In order to capture earnings variation in the criminal sector, the model allows for criminal earnings shocks in which the individual takes a new draw from $M(y)$.[10] If the individual accepts an income crime opportunity, the monthly criminal earnings remain constant within a spell until a criminal earnings shock arrives. The individual then chooses whether to exit income crime, or to continue with the new stream of criminal earnings. Besides the criminal earnings, individuals that participate in the criminal sector benefit from non-pecuniary benefits from income crime (α_c).[11]

The model incorporates the possibility of incarceration. An individual can be incarcerated in any state.[12] When incarcerated, the individual waits to be released at a given rate. The model also allows for the exogenous destruction of legal jobs and income crime.

The arrival rates of legal job offers (λ), income crime opportunities and criminal earnings shocks (η), incarceration (π), exogenous separations from income crime (τ), exogenous separations from legal jobs (δ), and releases from incarceration (κ) are all Poison processes that depend on the current state. The superscripts in the arrival rates index the current state.

The value of non-employment equals the flow utility of leisure (α_l), plus the expected value of changing labour market status. Non-employed individuals are subject to legal job offers drawn from $F(w)$ at rate λ^{ne}. If the individual accepts the offer, he transitions into legal employment. Individuals also face income crime opportunities drawn from $M(y)$ at rate η^{ne}. If the individual takes the opportunity, he transitions into income crime. Lastly, the individual is incarcerated and sent to jail at rate π^{ne}. The flow Bellman equation for a non-employed individual is

$$\begin{aligned}(r + \lambda^{ne} + \eta^{ne} + \pi^{ne})V^{ne} \quad &= \quad \alpha_l + \lambda^{ne} \int \max\left[V^e(x), V^{ne}\right] dF(x) \\ &\quad +\eta^{ne} \int \max\left[V^c(x), V^{ne}\right] dM(x) + \pi^{ne} J.\end{aligned} \tag{4.1}$$

The value of legal employment equals the corresponding flow utility of leisure (α_l) and legal earnings (w) plus the expected value of receiving a new legal job offer at rate λ^e, plus the expected value of receiving an income crime opportunity at rate η^e, plus the expected value of an exogenous termination at rate δ^e, plus the expected value of incarceration at rate π^e. Different from non-employed individuals, upon accepting an income crime opportunity, legal

[10]Allowing for criminal earnings shocks is one way to account for earnings changes in the sector. The model does not allow for on-the-crime search (i.e., analogous to on-the-job search) in the criminal sector, since it is not trivial to determine career changes in the data for the criminal sector, where there are no labour contracts, and there are usually no employers.

[11]Non-pecuniary benefits in the legal labour sector are normalized to zero.

[12]The incarceration rate is allowed to be positive for individuals who are not currently participating in income crime to account for incarcerations that occur as a consequence of past criminal activities or other criminal activities not described in the model.

workers must also decide whether to participate only in the criminal sector or to split their time between the criminal and legal labour sectors.[13] Legal workers also face an exogenous job-destruction rate. In this case, the individual transitions into non-employment. The flow Bellman equation for a legal worker who works at a firm offering legal earnings w is

$$\begin{aligned}(r+\lambda^e+\eta^e+\pi^e+\delta^e)V^e(w) &= (1-h_e)\alpha_l + h_e w + \delta^e V^{ne} + \lambda^e \int \max\left[V^e(x), V^e(w)\right] dF(x) \\ &+\eta^e \int \max\left[V^c(x), V^{ec}(w,x), V^e(w)\right] dM(x) + \pi^e J.\end{aligned} \tag{4.2}$$

The value of income crime equals the corresponding flow utility of leisure (α_l), plus pecuniary and non-pecuniary benefits from income crime ($y+\alpha_c$), plus the expected value of changing labour market status. Similar to legal workers, they face offers in the legal labour sector, an exogenous destruction rate, and an incarceration rate. Different from legal workers, individuals participating in income crime face criminal earnings shocks at rate η^c. In this case, the current level of criminal earnings is no longer available, and the individual has to choose between non-employment and the new stream of criminal earnings. The flow Bellman equation for an individual devoted to income crime with criminal earnings y is

$$\begin{aligned}(r+\lambda^c+\eta^c+\pi^c+\tau^c)V^c(y) &= (1-h_c)\alpha_l + h_c(y+\alpha_c) + \lambda^c \int \max\left[V^e(x), V^{ec}(x,y), V^c(y)\right] dF(x) \\ &+\eta^c \int \max\left[V^c(x), V^{ne}\right] dM(x) + \tau^c V^{ne} + \pi^c J.\end{aligned} \tag{4.3}$$

The value of participating simultaneously in the criminal and legal labour sectors equals the flow utility of legal earnings (w), plus pecuniary and non-pecuniary benefits from income crime ($y+\alpha_c$), plus the expected value of receiving a legal job offer at rate λ^{ec}, plus the expected value of facing a criminal earnings shock at rate η^{ec} , plus the expected value of an exogenous termination from income crime and legal jobs at rates τ^{ec} and δ^{ec}, respectively, plus the expected value of incarceration at rate π^{ec}. Different from individuals who only participate in income crime, individuals participating in both sectors do not have leisure time. In addition, they have four alternatives when they face a criminal earnings shock. Specifically, the individual has to decide between taking the new stream of criminal earnings, taking the new stream of criminal earnings and quitting the legal job, keeping the legal job and quitting crime, or exiting to non-employment. The flow Bellman equation for an individual participating in both sectors with legal earnings w and criminal earnings y is

[13] The legal job is full time regardless of whether the worker participates or not in income crime. Note that in the model, individuals are not allowed to work on a part-time basis in the legal labour sector.

$$
\begin{aligned}
(r+\lambda^{ec}+\eta^{ec}+\pi^{ec}+\delta^{ec}+\tau^{ec})V^{ec}(w,y) \;=\;& h_e w + h_c(y+\alpha_c) \\
&+\lambda^{ec}\int \max\left[V^{ec}(x,y), V^{e}(x), V^{ec}(w,y)\right] dF(x) \\
&+\eta^{ec}\int \max\left[V^{ec}(w,x), V^{c}(x), V^{e}(w), V^{ne}\right] dM(x) \\
&+\delta^{ec}\max[V^{c}(y), V^{ne}] + \tau^{ec}\max[V^{e}(w), V^{ne}] + \pi^{ec} J.
\end{aligned}
\tag{4.4}
$$

Lastly, the value of jail equals the flow utility of incarceration (α_j) plus the expected value of being released at rate κ and facing an immediate legal job offer or income crime opportunity with probabilities ρ_e and ρ_c, respectively. The flow Bellman equation for an incarcerated individual is

$$
(r+\kappa)J \;=\; \alpha_j + \kappa\left[(1-\rho_e-\rho_c)V^{ne} + \rho_e\int \max[V^{e}(x), V^{ne}]dF(x) + \rho_c\int \max[V^{c}(x), V^{ne}]dM(x)\right]. \tag{4.5}
$$

4.2.1 Analysis of model properties

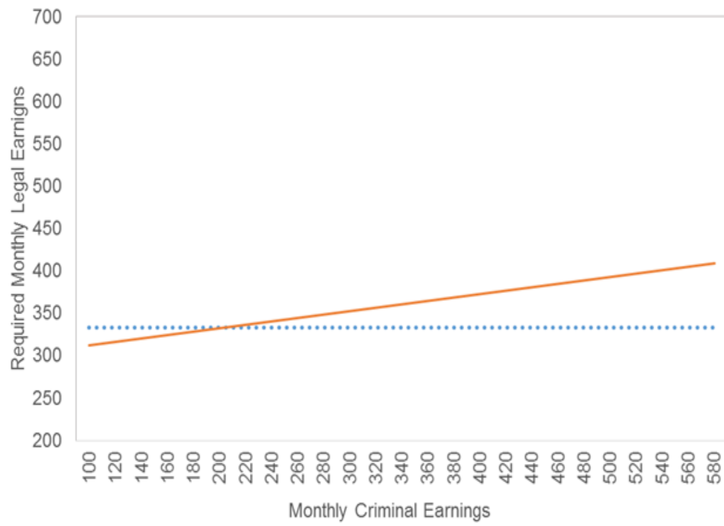
Individuals maximize future expected utility and decide whether to accept a legal job and participate in income crime by following a set of reservation rules. In this subsection, I define such reservation rules and discuss the main properties of the model.

I first consider the reservation legal earnings of non-employed individuals, and individuals participating in income crime. Since $V^e(w)$ is a continuous and increasing function in w and V^{ne} does not depend on w, a non-employed individual only accepts legal job offers that are at least as good as the reservation legal earnings denoted by w^*, and determined by $V^e(w^*) = V^{ne}$.[14] An individual engaged in income crime with criminal earnings y only accepts legal job offers that are at least as good as the reservation legal earnings denoted by $\bar{w}(y)$ and determined by $\max\{V^{ec}(\bar{w}, y),\ V^e(\bar{w})\} = V^c(y)$.[15] The reservation legal earnings of non-employed individuals and individuals participating in income crime are represented in Figure 4.1. The main conclusion that emerges from the figure is the positive dependence between criminal earnings and the reservation legal earnings for individuals participating in income crime. This implies that, the higher the criminal earnings, the higher the legal earnings required to accept a legal job offer. It is therefore important to distinguish between non-employment and income crime for the transitions across sectors. How individuals spend their time (e.g., participate or not in income crime) has implications for the employment choice in the legal labour sector. In a similar fashion, I can compare the reservation criminal earnings of non-employed individuals and individuals employed in the legal labour sector.[16]

[14] The non-employment reservation legal earnings (w^*) are also used for other choices: individuals who are released from jail and immediately receive a legal job offer and individuals participating in both sectors who are exogenously separated from crime.

[15] For individuals employed in the legal labour sector or participating in both sectors, the reservation legal earnings at which they are indifferent between accepting another legal job and staying with the current legal job

Figure 4.1: Reservation Legal Earnings for Individuals Participating in Income Crime



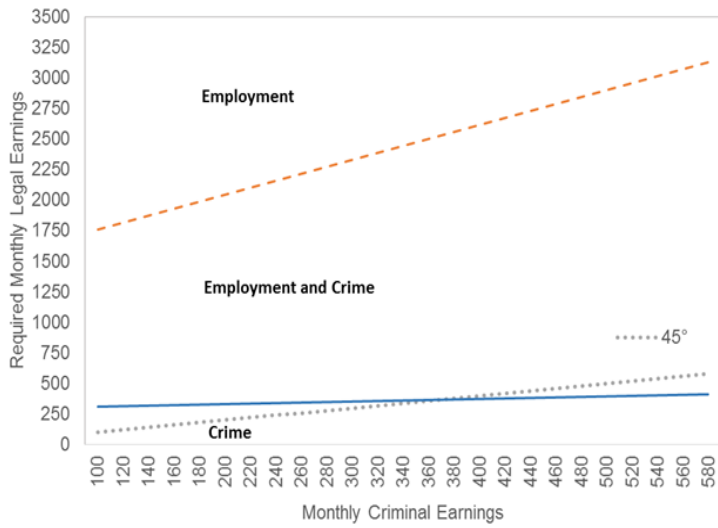
Notes:
1. The figures are based on the estimates from the baseline model in Table 4.3.
2. The reservation legal earnings curve (solid line) represents the minimum level of legal earnings required to accept the job offer by individuals participating in income crime.
3. As a benchmark, the reservation legal earnings for unemployed individuals is shown (dotted line).

The set of legal job offers accepted by individuals engaged in income crime can be further decomposed to account for the other decisions faced by them. Namely, individuals participating in income crime who accept a legal job offer, can either participate in both sectors or quit income crime. In Figure 4.2, the reservation legal earnings of individuals participating in income crime, as well as the reservation legal earnings that determines whether individuals participate in both sectors or quit income crime (second reservation legal earnings), are shown. The main conclusion is that individuals will not necessarily quit income crime when accepting a legal job offer. Individuals accept any offer above the reservation legal earnings, but they only quit income crime if the offer exceeds the second reservation legal earnings value. The set of income crime opportunities accepted by individuals employed in the legal labour sector can be decomposed in a similar way.

is their current level of legal earnings.

[16]Given that $V^c(y)$ is a continuous and increasing function in y and V^{ne} does not depend on y, an non-employed individual only accepts income crime opportunities that are at least as good as the reservation criminal earnings denoted by y^*, and determined by $V^c(y^*) = V^{ne}$. The non-employment reservation criminal earnings are also used by individuals who are released from jail that immediately receive an income crime opportunity, individuals participating in both sectors who are exogenously separated from the legal labour sector, and individuals participating in income crime that face a criminal earnings shock.

Figure 4.2: Decomposition of the Reservation Legal Earnings for Individuals Participating in Income Crime



Notes:
1. The figures are based on the estimates from the baseline model in Table 4.3.
2. The solid curve represents the minimum level of legal earnings required to accept the job offer. The dashed curve represents the minimum level of legal earnings required to accept the offer and quit crime.

4.2.2 Estimation

The set of parameters to estimate (θ) includes the mobility parameters (λ^{ne}, λ^{e}, λ^{c}, λ^{ec}, η^{ne}, η^{e}, η^{c}, η^{ec}, δ^{e}, δ^{ec}, τ^{c}, τ^{ec}, π^{ne}, π^{e}, π^{c}, π^{ec}, κ, ρ_e, and ρ_c), the flow utility parameters (α_l, α_c, and α_j), and the earnings distributions parameters (μ_w, μ_y, σ_w, and σ_y). The monthly interest rate is set to 0.4%, yielding a real annual interest rate of 5%.

The model is estimated using the spell data from the Pathways to Desistance described in Chapter 3. It is estimated separately for each location via Indirect Inference (Gourieroux, Monfort, and Renault, 1993).[17] The idea behind this method is to find a set of structural parameters that minimize the distance between a set of moments from the real data and the data simulated using the model and the values of the parameters. The moments used for the estimation help to identify the parameters and capture the main features of the model.

I assume that the legal and criminal earnings distributions are log normal. The parameters of the earnings distribution $F(w)$ are then identified from the accepted legal earnings information. Similarly, the parameters of the criminal earnings distribution $M(y)$ are identified from data on accepted criminal earnings. Hence, I use the first and second moments of accepted earnings in each sector for the estimation. Since there is considerable missing criminal earnings

[17]Descriptive statistics by location are reported in Table B.1 in Appendix B. The main features of the criminal and legal labour sectors documented for the overall sample (e.g., earnings premium in the criminal sector relative to the legal labour sector) are preserved in both locations.

information, I add a missing at random process for criminal earnings.[18] Specifically, the crime spell starts with missing earnings with probability p, and the arrival rates nm and m determine the switching rate from missing to non-missing and from non-missing to missing, respectively. These parameters are identified by the share of crime spells starting without criminal earnings information, the duration of missing earnings spells, and the duration of non-missing criminal earnings spells. Since the missing process is assumed to be random, it does not affect the identification strategy of the parameters of the criminal earnings distribution.

I follow Kiefer and Neumann (1993) and use the minimum legal earnings and criminal earnings in the data as the superefficient estimators of the reservation legal earnings and the reservation criminal earnings for the non-employed, respectively (i.e., $w^* = \min(w)$ and $y^* = \min(y)$).[19] Based on these values, I can identify the flow utility associated with leisure (α_l) and income crime (α_c), using the formula for the reservation earnings for the non-employed.[20]

The mobility parameters can be identified by durations and transition information. Therefore, I include moments concerning the durations of each state as well as conditional transitions by state to identify mobility parameters. There are eight sets of mobility parameters in the model which are allowed to vary by state: job arrival rates (λ^{ne}, λ^{e}, λ^{c}, λ^{ec}), arrival rates of income crime opportunities and criminal earnings shocks (η^{ne}, η^{e}, η^{c}, η^{ec}), income crime separation rates (τ^{c}, τ^{ec}), job destruction rates (δ^{e}, δ^{ec}), incarceration rates (π^{ne}, π^{e}, π^{c}, π^{ec}), jail release rate (κ), and the probabilities of immediate offers after jail (ρ_e and ρ_c). The model dictates that the transition probability between any two states is equal to the corresponding arrival rate times the probability that the individual chooses to make the transition. Flinn and Heckman (1982) show that transition information is enough to identify the mobility parameters as long as the earnings offer distribution is assumed to be recoverable. Intuitively, once we know the distribution of earnings and the minimum earnings accepted by individuals, the transition probabilities can be used to identify arrival rates. For example, non-employment-to-legal employment transitions identify the arrival rate of legal jobs for the non-employed. Likewise, non-employment-to-crime transitions and non-employment-to-jail transitions identify the arrival rate of income crime opportunities and the incarceration rate, respectively, for the non-employed. The remaining arrival rates of legal jobs, income crime, incarceration and the destruction rates are identified analogously, using the corresponding transitions. The arrival rate of criminal earnings shocks are identified using crime-to-crime transitions in which the average monthly earnings change by more than 10%. Lastly, the transitions from jail identify the

[18] Around 43% of the individual-month observations where the individual participates in crime has missing earnings. An alternative is to treat these observations as periods of participation in income crime but with zero criminal earnings, instead of missing criminal earnings. As a robustness check, I estimated the model using this alternative approach to dealing with observations with no reported criminal earnings. Specifically, upon facing a crime opportunity or a criminal earnings shock, there is a probability that the criminal earnings offered are zero. In this specification, the minimum criminal earnings accepted are zero. The results were very similar to my baseline estimates in Section 4.3.

[19] As is standard in search models, I assume that individuals are not willing to accept negative earnings to participate in either sector. If this assumption does not hold for this population, I am imposing an upper bound on α_l and α_c. The implied minimum monthly earnings are \$102 per month in the criminal sector, and \$322 per month in the legal labour sector.

[20] The reservation legal earnings for the non-employed are determined by $V^e(w^*) = V^{ne}$. Likewise, the reservation criminal earnings for the non-employed are determined by $V^c(y^*) = V^{ne}$.

release rate, as well as the probabilities of obtaining an immediate offer in either sector.

The flow utility associated with being incarcerated (α_j) is identified using two additional moments. I use the share of months that individuals participate in income crime and the share of months that individuals participate in both sectors as a share of the total months spent on income crime. The intuition is that a larger value of α_j should make the criminal sector more attractive, and it would have differential effects on income crime and employment/crime depending on the respective incarceration rates. The flow utility of being incarcerated is separately identified from the flow utility of income crime, α_c, since a shift in the latter only affects the criminal sector whereas a shift in α_j affects any state, as long as the incarceration rates are nonzero and they differ for individuals participating in crime depending on their participation in the legal labour sector. The full list of moments can be seen in Table 4.1.

The estimation procedure works as follows. First, I estimate the parameters of the missing process for criminal earnings and keep them fixed throughout the indirect inference procedure. Second, I guess values of all parameters, except for α_l and α_c. For a given guess of parameters, I solve for the value functions using fixed point methods and I obtain the implied values of the flow utility of leisure and income crime (α_l and α_c) as described above. Next, I simulate data based on these parameters. For the data simulation, I mimic the sampling scheme of the original data. In particular, I draw a vector of pseudo-random draws that determine the initial state and initial survey.[21] I also draw a vector of pseudo-random draws that determine the probability of attrition conditional on the survey (initial and posterior surveys).[22] From this simulated data, I calculate the set of selected moments. The indirect inference estimate of the structural parameters minimizes the difference between the simulated and the corresponding moments from the data. Let g represent the vector of moments in the data and let $g(\theta)$ represent the vector of simulated moments given the parameter values θ. The criterion function is then,

$$\Phi(\theta) = (g - g(\theta))' W^{-1} (g - g(\theta))$$

where W is a weighting matrix. I use a diagonal weighting matrix during estimation, where each diagonal element is the variance of the corresponding moment. I calculate the matrix W by bootstrapping 500 samples from the original sample of data and calculating the sample moments for each bootstrapped sample. I minimize the objective function using Gauss-Newton.

4.3 Results

I now present the results from my baseline specification separately for each location. Before discussing the parameter estimates, I judge the fit of the model by examining the moments I ex-

[21]By design, each individual in the PDS completes at most ten surveys. Since the sample used for estimation concentrates on individuals once they have transitioned out of school, the initial survey in the sample can be any survey between the first and the tenth.

[22]Individuals drop out of the sample for two main reasons. First, they are not interviewed again after completing the tenth survey. Second, they can voluntarily drop out (i.e., attrition). In the simulated data, individuals are not observed beyond the tenth survey and they face a probability of attrition at any point in time.

Table 4.1: Model Fit - Data and Estimated Moments by Location

Moment	Philadelphia Data	Philadelphia Model	Phoenix Data	Phoenix Model
Earnings				
Average Ln(Monthly Legal Earnings)	6.989	6.989	7.084	7.083
Standard Deviation Ln(Monthly Legal Earnings)	0.469	0.470	0.434	0.434
Average Ln(Monthly Criminal Earnings)	7.963	7.967	7.060	7.062
Standard Deviation Ln(Monthly Criminal Earnings)	1.150	1.154	1.290	1.289
Adjusted Durations				
Non-Employment Duration	7.503	7.441	5.408	5.398
Legal Employment Duration	6.460	6.393	7.136	7.120
Incarceration Duration	14.353	14.325	12.909	12.849
Income Crime Duration	4.731	4.783	3.392	3.417
Employment/Crime Duration	3.963	3.713	3.603	3.578
Conditional Transitions				
Non-Employment to				
Legal Employment	0.499	0.495	0.550	0.548
Income Crime	0.256	0.261	0.200	0.203
Incarceration	0.245	0.244	0.250	0.249
Legal Employment (E) to				
Non-Employment	0.554	0.559	0.371	0.378
New Legal Job	0.200	0.206	0.335	0.344
Incarceration	0.068	0.048	0.042	0.012
Income Crime	0.049	0.049	0.065	0.067
Employment/Crime	0.129	0.138	0.187	0.199
Income Crime to				
Non-Employment	0.342	0.350	0.363	0.382
Legal Employment	0.039	0.009	0.058	0.006
Incarceration	0.333	0.339	0.300	0.314
Income Crime	0.234	0.245	0.158	0.170
Employment/Crime	0.052	0.057	0.121	0.128
Employment/Crime (EC) to				
Non-Employment	0.044	0.027	0.020	0.020
Same Legal Job (E)	0.441	0.447	0.518	0.518
Same Legal Job (EC)	0.103	0.108	0.161	0.161
New Legal Job (E or EC)	0.059	0.058	0.047	0.046
Income Crime	0.235	0.235	0.202	0.202
Incarceration	0.118	0.125	0.052	0.053
Incarceration to				
Non-Employment	0.753	0.749	0.576	0.575
Legal Employment	0.079	0.080	0.228	0.227
Income Crime	0.168	0.171	0.196	0.198
Additional Moments				
Share of Months Income Crime	0.188	0.184	0.161	0.158
Share of Employment/Crime	0.220	0.211	0.448	0.450

Notes:

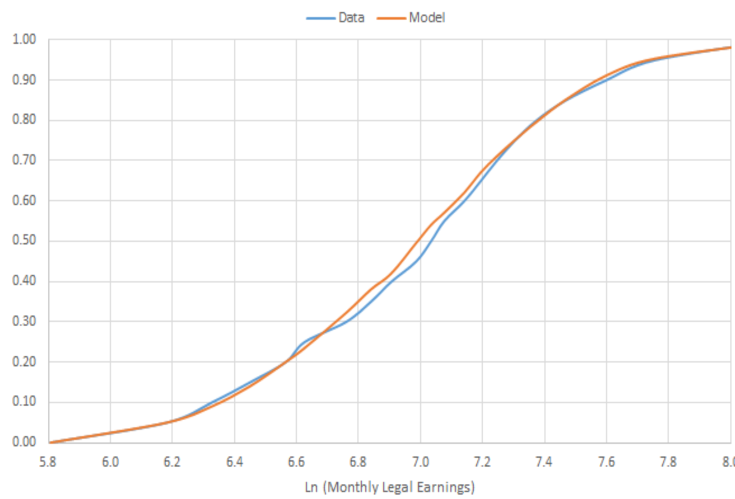
1. Adjusted durations are in months and are calculated as the sum of durations of censored and uncensored spells, over the number of uncensored spells. Income crime duration is the duration of income crime spells ignoring crime-to-crime transitions. Likewise, employment/crime duration equals the duration of employment/crime spells, ignoring employment/crime-to-employment/crime transitions.

3. Conditional transitions refer to spell to spell transitions, for uncensored spells. The transition probabilities sum to one since I only consider completed (i.e., uncensored) spells for the calculation.

4. Non-employment-to-employment/crime and incarceration-to-employment/crime transitions represent less than 1% of all transitions corresponding to non-employment and incarceration spells. As a result, they are coded as missing transitions in the data. In the model, these are zero probability events. Income crime-to-income crime transitions entail at least a 10% change in monthly criminal earnings, conditional on participating in income crime.

plicitly target in the estimation procedure. The sample and estimated moments are reported in Table 4.1 for Philadelphia and Phoenix. The model does a very good job in fitting the first and second moments of the accepted earnings distributions in the criminal and legal labour sectors across the two locations.[23] The model also fits well average durations and conditional spell transitions, with only a few modest discrepancies for some of the lower probability transitions, such as the crime-to-employment transition.[24] I also obtain a good fit for the share of months engaged in income crime.

Figure 4.3: Model Fit - Cumulative Distribution Function - Ln (Monthly Legal Earnings) - Philadelphia



Notes:
1. The figure is based on the estimates from the baseline model in Table 4.3.

Parameter estimates for the baseline model are presented in Table 4.3 for Philadelphia and Phoenix.[25] I focus on the results for Philadelphia and discuss the results for Phoenix only when there are important differences in the parameter estimates across locations.

The results reflect the earnings premium in the criminal sector relative to the legal labour sector.[26] However, the variance of earnings in the criminal sector is larger as well. The higher

[23]In Figures 4.3, 4.4, 4.5, and 4.6, I show that the model fits well the cumulative distribution function of the logarithm of monthly earnings in the legal labour and criminal sectors, respectively, in the two locations.

[24]The results in Table 4.2 indicate that the model also fits well the unconditional transitions (month to month transitions), particularly for non-employment, legal employment, and incarceration.

[25]As a robustness check, I estimated the model using different trimming percentages for earnings in the criminal sector (e.g., 1% or 5% in the top and bottom of the distribution). The results are shown in Tables C.1 and C.2 in Appendix C. Most of the parameters remain practically unchanged, except for the parameters of the criminal earnings distribution, and the flow utility of jail and crime, given that the minimum, average, and standard deviation of accepted criminal earnings used to identify them are directly affected.

[26]The parameters from the missing process for criminal earnings indicate that 53% of the income crime spells start with missing earnings information. The switching rates from missing to nonmissing and from nonmissing to missing are estimated at 0.36 and 0.20, respectively.

Table 4.2: Additional Data and Estimated Moments - Unconditional Transitions by Location

Moment	Philadelphia		Phoenix	
	Data	Model	Data	Model
Non-Employment to				
Same Spell	0.870	0.855	0.816	0.817
Legal Employment	0.065	0.072	0.101	0.101
Income Crime	0.033	0.035	0.037	0.045
Incarceration	0.032	0.038	0.046	0.037
Legal Employment to				
Same Spell	0.848	0.840	0.863	0.859
Non-Employment	0.084	0.089	0.051	0.054
Legal Employment	0.030	0.033	0.046	0.048
Incarceration	0.007	0.008	0.009	0.009
Income Crime	0.010	0.008	0.006	0.002
Employment/Crime	0.020	0.022	0.026	0.028
Income Crime to				
Same Spell	0.709	0.748	0.642	0.691
Non-Employment	0.099	0.088	0.130	0.119
Legal Employment	0.011	0.001	0.021	0.000
Incarceration	0.097	0.086	0.108	0.099
Income Crime	0.068	0.061	0.056	0.053
Employment/Crime	0.015	0.015	0.043	0.039
Employment/Crime to				
Same Spell	0.706	0.733	0.658	0.710
Non-Employment	0.013	0.006	0.007	0.005
Legal Employment	0.147	0.120	0.193	0.149
Income Crime	0.035	0.034	0.018	0.015
Incarceration	0.069	0.063	0.069	0.059
Employment/Crime	0.030	0.044	0.055	0.060
Incarceration to				
Same Spell	0.928	0.910	0.943	0.908
Non-Employment	0.054	0.067	0.033	0.053
Legal Employment	0.006	0.007	0.013	0.021
Income Crime	0.012	0.015	0.011	0.018

Notes:

1. Unconditional transitions refer to month-to-month transitions, including censored and uncensored spells.

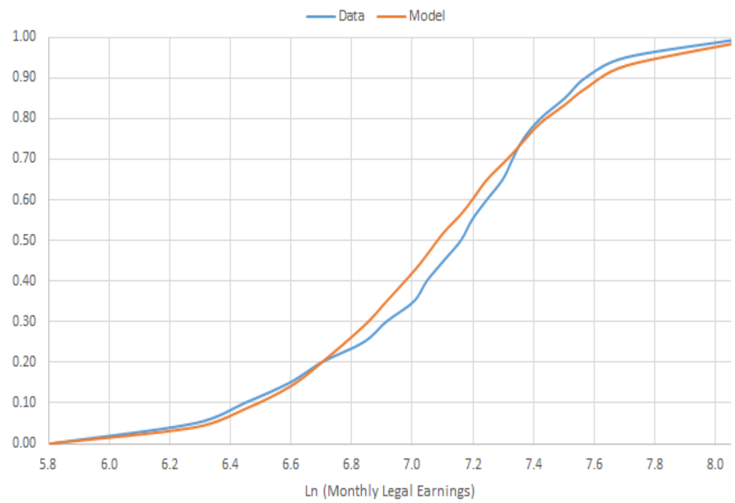
2. Non-employment-to-employment/crime and incarceration-to-employment/crime transitions represent less than 1% of all transitions corresponding to non-employment and incarceration spells. As a result, they are coded as missing transitions in the data. In the model, these are zero probability events.

Table 4.3: Parameter Estimates by Location

Parameter	**Philadelphia**	**Phoenix**
Legal Earnings Distribution: Mean	7.281 (0.037)	7.339 (0.024)
Legal Earnings Distribution: Variance	0.240 (0.025)	0.200 (0.015)
Criminal Earnings Distribution: Mean	8.327 (0.183)	7.234 (0.301)
Criminal Earnings Distribution: Variance	1.401 (0.283)	2.114 (0.636)
Incarceration: Release Rate	0.075 (0.006)	0.081 (0.008)
Non-Employment: Job Arrival Rate	0.069 (0.007)	0.102 (0.008)
Non-Employment: Crime Arrival Rate	0.037 (0.005)	0.040 (0.007)
Non-Employment: Incarceration Rate	0.034 (0.004)	0.046 (0.005)
Legal Employment: Job Arrival Rate	0.082 (0.015)	0.141 (0.013)
Legal Employment: Crime Arrival Rate	0.031 (0.008)	0.031 (0.006)
Legal Employment: Job Destruction Rate	0.089 (0.009)	0.054 (0.004)
Legal Employment: Incarceration Rate	0.008 (0.002)	0.010 (0.002)
Income Crime: Job Arrival Rate	0.027 (0.010)	0.054 (0.012)
Income Crime: Earnings Shock Rate	0.123 (0.021)	0.113 (0.025)
Income Crime: Destruction Rate	0.101 (0.013)	0.132 (0.025)
Income Crime: Incarceration Rate	0.095 (0.011)	0.109 (0.014)
Employment/Crime: Job Arrival Rate	0.061 (0.036)	0.047 (0.015)
Employment/Crime: Crime Earnings Shock Rate	0.063 (0.031)	0.102 (0.019)
Employment/Crime: Crime Destruction Rate	0.130 (0.038)	0.172 (0.034)
Employment/Crime: Incarceration Rate	0.039 (0.014)	0.018 (0.006)
Employment/Crime: Job Destruction Rate	0.074 (0.027)	0.071 (0.010)
Incarceration: Job Offer Probability	0.086 (0.022)	0.234 (0.036)
Incarceration: Crime Opportunity Probability	0.165 (0.029)	0.208 (0.048)
Flow Utility of Incarceration	172.2 (261.7)	373.6 (317.4)
Flow Utility of Leisure	956.1 (431.8)	1433.8 (326.0)
Flow Utility of Crime	674.6 (1005.6)	1539.3 (581.1)

Notes:
1. Standard errors are reported below the point estimates in parenthesis. These are computed by bootstrap with 100 replications.
2. Arrival rates are monthly.
3. The flow utility of crime equals ($\alpha_c * 2/3$), which is the non-pecuniary value of crime obtained by an individual who is only participating in income crime.

Figure 4.4: Model Fit - Cumulative Distribution Function - Ln (Monthly Legal Earnings) - Phoenix



Notes:
1. The figure is based on the estimates from the baseline model in Table 4.3.

average offer makes the criminal sector more attractive relative to the legal labour sector, but the higher variance makes the offer distribution much more dispersed, increasing the likelihood of facing a low income draw that may be below the reservation value. While there are not essential differences in the distribution of legal earnings across locations, the distribution of criminal earnings in Philadelphia has first-order dominance over the distribution of criminal earnings in Phoenix. The gap between average earnings offered in the criminal sector and the legal labour sector is consequently larger in Philadelphia than in Phoenix.

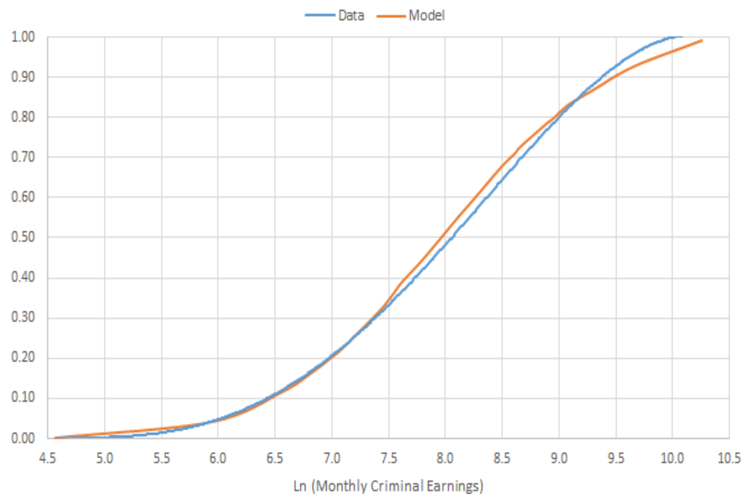
Besides earnings, I also find that the non-pecuniary benefits from crime are large relative to the non-pecuniary benefits from legal jobs, which are normalized to zero. The estimated value of non-pecuniary benefits from crime are $674 per month. Nevertheless, it is not precisely estimated.

Regarding the arrival rates of legal jobs, I find that it takes on average 14.5 months for non-employed individuals to receive an offer. The estimated arrival rate for non-employed individuals is twice as long as is usually estimated in the literature for similar demographic groups, suggesting that search frictions may be more prevalent for young serious offenders relative to the rest of the population.[27] One explanation is that serious offenders face certain employment restrictions in the legal labour sector, which directly limits their possibilities.[28]

[27]Using the NLSY79, Engelhardt (2010) concentrates on individuals with and without criminal records and estimates that they receive offers every 6.7 and 1.8 months, respectively when they are unemployed.

[28]Bushway and Sweeten (2007) document that ex-felons are barred from up to 800 different occupations across the United States.

Figure 4.5: Model Fit - Cumulative Distribution Function - Ln (Monthly Criminal Earnings) - Philadelphia



Notes:
1. The figure is based on the estimates from the baseline model in Table 4.3.

Another possibility is that firms restrain from hiring individuals with a criminal past.[29]

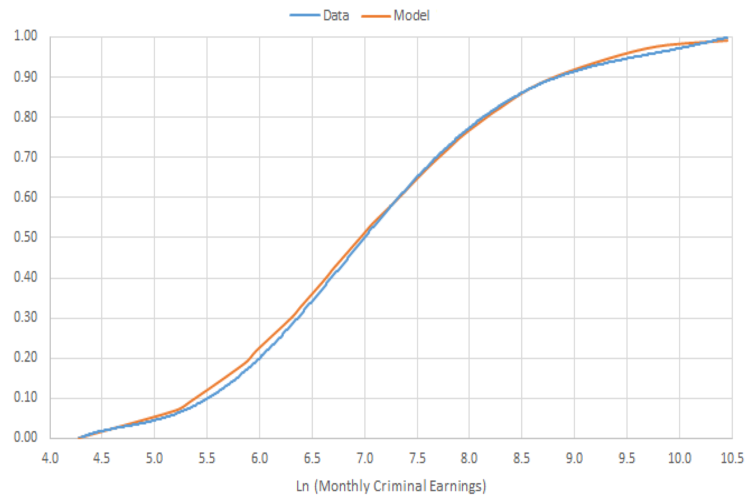
The arrival rate of legal job offers for legally employed individuals is larger, although not significantly different, than that of non-employed individuals. This is different from the standard search literature, where the job arrival rate of employed individuals is usually smaller than that of unemployed individuals. This suggests that participating in the legal labour sector may provide further opportunities to increase earnings through changing jobs within the legal labour sector. It is possible that for individuals with a criminal history, having a legal job is a good signal to get further offers, relative to being non-employed.

Individuals also receive legal job offers when they are participating in the criminal sector. I find that the arrival rate of legal job offers for individuals who are currently engaged in income crime is roughly half as large as the arrival rate for non-employed individuals, and one third as large as the arrival rate for individuals employed exclusively in the legal labour sector. The arrival rate of legal job offers for individuals participating in both sectors is also significantly smaller than that of individuals participating only in the legal labour sector.[30] One interpretation is that individuals participating in income crime have less time available and/or less incentives to search for jobs in the legal labour sector. Since in the model one way to receive higher earnings is by moving to another job in the legal labour sector, the overall lower arrival

[29]See the review by Schmitt and Warner (2011) for evidence on employers disfavoring applicants with records. One major obstacle that offenders face when trying to find legal employment is the question about their criminal history that often appears on applications for legal jobs (Agan and Starr, 2017; Pager, 2003; Uggen et al., 2014).

[30]This differential arrival rate is one distinctive feature of the model. Differently, in the model developed by Burdett, Lagos, and Wright (2003, 2004), crime has no direct implications on the job search intensity.

Figure 4.6: Model Fit - Cumulative Distribution Function - Ln (Monthly Criminal Earnings) - Phoenix



Notes:
1. The figure is based on the estimates from the baseline model in Table 4.3.

rates of legal jobs when participating in income crime partly deter people from participating in income crime. Furthermore, the larger arrival rate of legal job offers for individuals that already have a legal job further encourages participation in the legal labour sector.

Phoenix has, on average, larger arrival rates of legal jobs than Philadelphia. For example, it takes on average 9.8 months for non-employed individuals to receive a legal job offer in Phoenix, 4.7 months less than in Philadelphia. Only the arrival rate of legal jobs for individuals participating in the two sectors is larger in Philadelphia than in Phoenix, but these estimates have large standard errors.

Besides the high prevalence of search frictions, the legal labour sector for young offenders is also characterized by large destruction rates relative to estimates for a more general population (Eckstein and Van den Berg, 2007). The estimated destruction rate of legal jobs is 0.09 and 0.07 for individuals who participate and do not participate in income crime, respectively, in Philadelphia. The estimated job destruction rates are also large in Phoenix, although they are smaller than in Philadelphia. In Phoenix, the estimated destruction rate of legal jobs is 0.05 and 0.07 for individuals who participate and do not participate in income crime, respectively. One explanation for the large destruction rates is that the legal jobs available to previous offenders are of low quality and do not tend to last (e.g., temporary legal jobs). Another possibility is that individuals with a criminal background are fired more often than average workers because of bad behaviour. The large destruction rates in the legal labour sector can also be explained by the inability to exercise self-control, which at the same time can explain a large part of criminal behaviour (Gottfredson and Hirschi, 1990). Individuals also face large destruction rates in the criminal sector. The estimated destruction rates in the criminal sector are, on average, twice as large as those in the legal labour sector, making the criminal sector inferior to the legal labour sector in this aspect.

Income crime opportunities are also subject to frictions. The estimated arrival rate of income crime opportunities is significantly smaller than the arrival rate of legal jobs for non-employed individuals. Income crime opportunities arrive at a lower rate for individuals participating in the legal labour sector, but the arrival rates are not significantly different from each other. Lastly, individuals participating in income crime face criminal earnings shocks approximately every eight months, indicating that the criminal sector is quite volatile despite the high average earnings offered.

The criminal sector features high incarceration rates, making the criminal sector less attractive than legal employment. Not surprisingly, the incarceration rate for individuals who participate simultaneously in the criminal and legal labour sectors is significantly lower, relative to individuals fully devoted to the criminal sector.[31] When incarcerated, inmates have an expected incarceration time of 13.3 months. Individuals face an 8.6% probability of having a legal job offer immediately after being released, which reduces the cost of jail. These immediate offers may be rationalized by placement programs post-incarceration. Released individuals also face a 16.5% probability of getting an immediate income crime opportunity after jail. Consistent with the relative differences in the job arrival rates across locations commented above, individuals in Phoenix face a larger probability of a legal job offer post-jail (23.4%). Lastly, the flow utility of leisure is estimated at $956 per month. The flow utility of jail is $172 per month, although it is not precisely estimated.

All in all, the results show that disadvantaged young individuals face two poor-quality employment alternatives. On the one hand, the legal labour sector in the two locations is characterized by low arrival rates, low wages, and high destruction rates. The low arrival rates of legal job offers delay the transition from non-employment or crime to the legal labour sector, and generate long periods without legal earnings. The legal jobs offered to these individuals do not last long, generating quick transitions back to non-employment. On the other hand, the criminal sector appears as an alternative, both to non-employment and legal employment. The criminal sector offers an evident earnings premium relative to the legal labour sector, and has large non-pecuniary benefits. Nonetheless, like the legal labour sector, the criminal sector is characterized by low arrival rates and large destruction rates. In addition, the risk of being incarcerated and sent to jail reduces the value of crime. Lastly, there are a few important differences in the point estimates across locations. Search frictions in the legal labour sector are overall more prevalent in Philadelphia than in Phoenix, presenting lower job arrival rates and larger job destruction rates. Furthermore, the criminal earnings premium is larger in Philadelphia than in Phoenix, making the criminal sector even more attractive than the legal labour sector in this aspect. These estimated differences are used in the next section to illustrate how the criminal and legal labour sectors interact with each other.

[31] The incarceration rate for legally employed individuals is small and significantly different from that of individuals participating in crime (exclusively or not). The incarceration rate for non-employed individuals is smaller, but not significantly different, from that of individuals participating in crime and legal employment. As was mentioned in footnote 16, this does not seem to be driven by participation in non-income crimes.

4.4 Counterfactuals

In this section, I attempt to disentangle the roles of some features of the legal labour sector (e.g., job arrival rate) and the criminal sector (e.g., incarceration rate), in driving crime and legal employment choices. Understanding the role of these factors can shed light on alternative instruments available to policy makers interested in boosting legal employment or reducing youth crime. To that end, I present two sets of simulations. I start by illustrating the interactions between the criminal and legal labour sectors in Subsection 4.4.1. I then perform a policy analysis which shows that, as a result of the trade-offs between sectors, policies in one sector have important implications in the other sector. For instance, youth crime can be reduced both via policies targeting the criminal sector and policies that improve outcomes in the legal labour sector.

4.4.1 Interactions across sectors

In this subsection, I illustrate how the criminal and legal labour sectors interact with each other. I focus on the role of earnings and frictions in determining transitions between non-employment, crime, legal employment, and jail. Distinguishing the role of each of these factors is important since their implications are potentially different. For instance, both a reduction in average earnings offered in the criminal sector and an increase in the arrival rate of legal job offers result in a larger legal employment rate. However, they have opposite effects on the share of non-employed individuals. Furthermore, the size of the effects may differ across factors.

For this purpose, I take advantage of the existing differences between Philadelphia and Phoenix regarding the characteristics of the criminal and legal labour sectors: arrival rates of legal jobs, destruction rates of legal jobs, and the earnings gap between the criminal and legal labour sectors. I use the parameter estimates from Philadelphia, and modify a specific set of parameters using the estimates from Phoenix. In other words, I make Philadelphia look like Phoenix in a specific dimension, and discuss its implications on crime, legal employment, non-employment, and the incarcerated population. I start by simulating an increase in the arrival rate of legal jobs and a reduction in the destruction rate of legal jobs. I then concentrate on the criminal sector by simulating a change in the distribution of criminal earnings.

The simulations are performed as follows. Given alternative values for some of the parameters, I compute new value functions using fixed point methods. I then simulate the path of a sample of individuals, for whom the initial distribution across states equals the steady state distribution in the baseline model. Lastly, I compare the distribution across states and average earnings at the new steady state with those from the steady state in the baseline model.

The results are presented in Table 4.4. The first column of the table shows the average monthly distribution of the population across different states of the model at the steady state, using the baseline parameter estimates for Philadelphia. The next columns show the changes in the predicted outcomes under alternative values of the parameters relative to the baseline parameter estimates. For this exercise, the monthly crime rate is defined as the sum of the share of individuals participating in income crime, regardless of their legal employment status. The monthly legal employment rate is defined as the sum of the share of individuals participating in legal employment, regardless of whether they participate in income crime.

In columns 2, 3, and 4 of Table 4.4, I show how non-employment, legal employment, crime, and the incarcerated population in Philadelphia are affected when changing some characteristics of the legal labour sector. I first simulate an increase in the arrival rates of legal jobs (λ^u, λ^e, λ^c, λ^{ec}) using the parameter estimates for Phoenix. This is equivalent to increasing the arrival rate of legal jobs by 44%, on average, or reducing by six months the average time elapsed until a legal job offer arrives.[32] Column 2 shows that facilitating access to legal jobs in Philadelphia results in a large increase in legal employment (7.8 percentage points or 31.0%). Because individuals find legal jobs at a faster rate, they spend, on average, less time non-employed, and the monthly crime rate goes down, but by only 0.4 percentage points (3.0%).[33] The shift in crime is modest, because the share of individuals participating simultaneously in the two sectors actually increases. Individuals participating in income crime benefit from the smaller search frictions, but they do not necessarily quit crime upon accepting the legal job. Average accepted earnings in the legal labour sector increase, and legal jobs have a smaller average durations due to the possibility of climbing up the ladder through job-to-job transitions.

Second, I simulate an increase of the average time before the legal job gets exogenously destroyed using the estimated job destruction rates in Phoenix. This is equivalent to an average reduction of 30% in the destruction rates of legal jobs (δ^e, δ^{ec}) in Philadelphia. The results are shown in column 3. The effects are similar to the previous simulation, generating a big increase in legal employment (8.3 percentage points or 32.8%) and a small reduction in crime (1.1 percentage points or 8.8%). One difference is that the average legal job duration increases, because reducing the job destruction rate directly reduces the flow of individuals from legal jobs to non-employment.

Next, I combine the last two simulations by simulating jointly an increase in the arrival rate and a reduction in the destruction rate of legal jobs. In other words, both the quality and the access to the legal labour sector in Philadelphia improve. The results in Column 4 show that there are complementarities in terms of crime reduction and legal employment boost when the two set of parameters are modified. These changes generate an increase in legal employment of 17.1 percentage points (90.0%) and a 1.9 percentage point reduction in crime (16.0%), exceeding the sum of the effects from the individual simulations in columns 2 and 3. Overall, increasing the arrival rate of legal jobs and reducing the job destruction rate have large effects on legal employment and a more modest, but non-negligible, effect on crime.

[32] Job placement programs are one example of policies that can reduce frictions in the legal labour sector, particularly for individuals with a criminal background (Wilson et al., 1999; Uggen, 2000).

[33] This result is consistent with prior empirical research that documents a reduction in criminal activity among previous offenders who face more legal employment opportunities (Uggen, 2000; Schnepel, 2016; Heller, 2014).

Table 4.4: Changes in the Parameter Estimates in Philadelphia

	(1) **Baseline Philadelphia**	**(2)** **Arrival rate of jobs in Phoenix**	**(3)** **Destruction rate of jobs in Phoenix**	**(4)** **Legal labour sector (2)+(3)**	**(5)** **Criminal earnings distribution parameters in Phoenix**
		Changes in %-points			
Non-Employment	33.96	-4.29	-4.39	-8.92	3.90
	[33.38;34.48]	[-4.95;-3.61]	[-5.03;-3.76]	[-9.56;-8.21]	[3.23;4.61]
Legal Employment (E + EC)	25.25	7.84	8.28	17.11	3.40
	[24.68;25.87]	[6.97;8.65]	[7.54;9.03]	[16.3;17.9]	[2.71;4.09]
Crime (C + EC)	12.04	-0.35	-1.06	-1.91	-4.65
	[11.69;12.4]	[-0.82;-0.1]	[-1.52;-0.64]	[-2.39;-1.41]	[-5.11;-4.23]
Incarceration	31.36	-2.32	-2.63	-5.41	-3.26
	[30.7;31.99]	[-3.11;-1.51]	[-3.37;-1.89]	[-6.21;-4.6]	[-3.95;-2.55]

Notes:

1. The first column shows the predicted outcomes of the model at the steady state using the baseline parameter estimates for Philadelphia from Table 11. In the next columns, I simulate changes in a specific set of parameters in Philadelphia. In column (2), I use the estimated arrival rates of legal jobs in Phoenix. In column (3), I use the estimated destruction rates of legal jobs in Phoenix. In column (4), I change both the arrival rate and destruction rate of legal jobs in Philadelphia using the parameter estimates for Phoenix. In column (5), I change the estimated parameters of the criminal earnings distribution using the parameter estimates for Phoenix. See Table 11 for the parameter estimates for Phoenix. In each scenario, the rest of the parameters estimates remain unchanged.
2. The figures in columns (2) to (5) represent the changes in the predicted outcomes in Philadelphia. All the changes are expressed in percentage points.
3. The numbers in brackets are the 95% bootstrap confidence intervals using 100 replications.

Next, I show how outcomes are affected by shifting parameters in the criminal sector. I focus specifically on the criminal earnings premium. Average earnings offered in the criminal sector in Philadelphia are more than four times larger than average earnings offered in the legal labour sector. The criminal sector in Phoenix also offers an earnings premium, although it is smaller: average earnings offered in the criminal sector are just two times larger than those offered in the legal labour sector. Overall, the distribution of criminal earnings in Philadelphia has first-order dominance over the criminal earnings distribution in Phoenix, and there are not important differences in the distribution of legal earnings across locations. In column 5, I simulate a reduction in earnings offered in the criminal sector in Philadelphia by using the estimates of the criminal earnings distribution for Phoenix. Reducing the criminal earning premium has a large effect on crime, with a reduction in the monthly crime rate of 4.7 percentage points (38.6%). It also generates a 3.4 percentage point increase in legal employment (13.5%). Unlike the rest of the simulations, the share of non-employed individuals increases. Due to the reduction in the criminal earnings premium, the relative value of the criminal sector decreases and, as a result, a larger share of individuals is now waiting in non-employment to get an attractive job offer in the legal labour sector rather than taking an opportunity in the criminal sector.

Overall, the results stress that there are sizable interactions between the criminal and legal labour sectors. I show that reducing the search frictions in the legal labour sector can both increase legal employment and reduce youth crime. The results also suggest that there is room for improvement in legal employment through changes in the criminal sector (e.g., reduction in criminal earnings premium).

4.4.2 Policy Simulations

In this section, I simulate the effect of three policies that target a one percentage point reduction in the monthly crime rate in Philadelphia.[34] While two are traditional policies that target the criminal sector, one policy targets the legal labour sector instead. As was suggested in Section 4.4.1, because of the sizable interactions across sectors, affecting the legal labour sector can also generate reductions in youth crime. Targeting the legal labour sector also has different implications on non-employment, legal employment, and jail, relative to policies that focus on the criminal sector.

I start by simulating a policy that extends the average sentence length. Second, I simulate an increase in the number of police. Lastly, I simulate the introduction of a wage subsidy in the legal labour sector. For each policy, the average monthly crime rate at the steady state using the corresponding values of the parameters is compared against the monthly crime rate at the steady state when using the baseline parameter estimates. For this exercise, I use the parameter estimates for Philadelphia.

I perform a back-of-the-envelope cost analysis to compare the efficiency of the policies.

[34]the monthly crime rate is defined as the sum of the share of individuals participating in income crime, regardless of their legal employment status.

The net cost of a policy includes the change in benefits paid to individuals who do not have a legal job (i.e., non-employed individuals and individuals participating exclusively in income crime) like welfare programs, the change in income taxes collected from legal workers, the change in prison expenditures, the change in costs due to income crime, and any direct costs specific to each policy. I assume that the benefits paid to individuals without legal jobs amount to \$600 per month.[35] To determine the prison expenditures, I use the average annual cost per inmate in Pennsylvania (Stephan, 2001). I assume a tax rate on legal earnings of 13.8%, which is approximately equal to the sum of the federal tax rate for low earnings and the state tax rate on income in Pennsylvania. To determine the cost of crime, I first calculate the unit cost of a crime, by type of income crime. For each type of income crime, I obtain the monthly cost by multiplying the average number of that type of crime committed in a given month in the PDS and the corresponding cost of crime.[36] The total monthly cost of income crime is then calculated as the the sum of the monthly cost of each type of crime. I use the estimates of the intangible costs of drug-related crimes from Rajkumar and French (1997). For non drug-related income crimes, I use the victim costs and property losses in Miller, Cohen, and Wiersema (1996).[37] The total estimated monthly cost of crime amounts to \$1,541 per month per individual exclusively engaged in income crime.[38] The results provide a rough estimate of the cost of each policy, which permits comparisons across them. The results of the policies are summarized in Table 4.5).

The literature has found strong linkages between longer sentence lengths and lower crime rates (Durlauf and Nagin, 2010). I now simulate an extension of the average sentence length by lowering the release rate (κ). In order to achieve a one percentage point reduction in the monthly crime rate in Philadelphia, the release rate has to be 17.5% smaller, which is equivalent to extending the average sentence length by roughly 3 months.[39] As it is shown in the second column of Table 4.5, this produces a large increase in the incarcerated population (3.6 percentage points or 11.0%). The share of non-employed individuals goes down, while legal employment stays relatively flat. I further decompose the total effect on crime into an incapacitation effect and a deterrence effect. The former captures the effect of keeping the individuals off the streets, whereas the latter reflects the discouraging effect of a harsher punishment. I distinguish between these two effects by separately calculating the probability of being out of prison, and the probability of incarceration conditional on not being already in prison. I find that incapacitation represents 43.4% of the total effect of the harsher sentence length on crime,

[35] The underlying assumption is that α_l is composed of 50% benefits paid to non-employed individuals and that the rest is the utility of leisure (Engelhardt, Rocheteau, and Rupert, 2008).

[36] The average monthly frequency of income crime in PDS is 2.7 burglaries, 2.2 larcenies, 0.4 motor vehicle thefts, 18.7 drug-related crimes, and 0.3 robberies. The victim cost and property loss per crime amounts to \$624 for burglary, \$154 for larceny, \$1,060 for motor vehicle theft, and \$1,680 for robbery (Miller, Cohen, and Wiersema, 1996). The intangible cost of a drug-related crime is estimated at \$3 (Rajkumar and French, 1997). All the amounts are expressed in 2000 dollars.

[37] The victim cost is an estimate of productivity and wage losses, medical costs, and quality of life reductions. I assume that the total cost of a single crime comprises the victim cost and 20% of the property loss, since property losses are usually considered a partial transfer from the victim to the criminal.

[38] The monthly cost of income crime is assumed to be half of it if the individual participates simultaneously in the criminal and legal labour sector.

[39] The elasticity of crime with respect to the average sentence length is -0.35, which is within the range typically found in the literature (Levitt, 2004).

which partly explains the large increase in the incarcerated population. The distinction between incapacitation and deterrence is key to determining the cost associated with changes in the sentence length (Kessler and Levitt, 1999; Durlauf and Nagin, 2010). The net annual cost of this policy amounts to $748 per individual in the model.

I next simulate an increase in the number of police, which has been largely recognized in the literature as an effective tool to reduce crime.[40] Increasing the number of police can have two main effects: increase the incarceration rate and/or reduce the arrival rate of crime opportunities.[41] I simulate these two effects separately, assuming in each case that the desired reduction in crime is obtained entirely by affecting either the incarceration rate or the arrival rate of crime opportunities. Since police does not enter explicitly in the model, I use the elasticity of crime with respect to the number of police reported by Levitt (2004) to calculate the necessary change in the number of police officers. I then use the mean annual wage of police officers in Pennsylvania (Bureau of labour Statistics) to assess the direct cost of the policy.

I start by simulating a proportional increase in the incarceration rates (π^{ne}, π^{e}, π^{c}, π^{ec}) via an increase the number of police in Philadelphia. In order to obtain a one percentage point reduction in crime, the incarceration rates in Philadelphia have to increase by 13.1%. As it is shown in Column 3, increasing the incarceration rates yields a large increase in the share of individuals incarcerated. One difference of this policy, relative to extending the average sentence length, is that 34.3% of the total effect on crime comes through incapacitation. The incapacitation effect explains less of the total effect, and as a result, this policy generates a smaller increase in incarcerated population. The net annual cost of this policy is consequently smaller than a policy that extends the average sentence length; it amounts to $471 per individual in the model.

I next simulate a proportional reduction in the arrival rates of income crime opportunities (η^{ne}, η^{e}) via an increase in the number of police. In this case, I assume that the effect of having more police comes exclusively through the arrival rate of crime opportunities, and that it has no effect on the incarceration rates. The results in Column 4 show that the arrival rates of income crime opportunities have to decrease by 10.7%, on average, to achieve the desired reduction in crime. The greater search frictions in the criminal sector do not generate an increase in the incarcerated population, like traditional policies targeting the criminal sector do (e.g., higher incarceration rate or longer sentence length). The implications on the share of non-employed and legally employed individuals are also different. They both increase, although none of these effects is precisely estimated. The net annual cost of the policy amounts to $-218 per individual in the model. Even if by construction the direct cost of the policy (i.e. cost of extra police officers) is the same as the increase in the incarceration rate, the different implications on the incarcerated population, non-employment, and legal employment explain the much smaller cost.

[40] Some papers that review the literature on the effect of police on crime include Cameron (1988); Nagin (1998); Levitt (2004), among others.

[41] This issue relates to the literature considering different policing strategies like the adoption of "hot spots" policing, "problem-oriented" policing, among other strategies (Braga, 2001, 2005; Braga, Papachristos, and Hureau, 2014).

Table 4.5: Alternative Policies to Achieve a 1 Percentage Point Decrease in the Monthly Crime Rate in Philadelphia

	(1) Longer Average Sentence Length	(2) Higher Incarceration Rate	(3) Fewer Crime Opportunities	(4) Wage Subsidy
Change in the corresponding parameter estimate	-17.41%	13.09%	-10.65%	619.52
	[-20.73;-14.30]	[10.37;15.33]	[-13.39;-8.16]	[504.80;746.35]
Annual Net Cost per individual (2000 USD)	748.62	471.62	-218.79	1327.44
	[545.55;929.11]	[329.38;598.21]	[-360.96;-89.25]	[969.80;1686.47]
Change in share non-employed (%-points)	-1.94	-0.77	0.72	0.35
	[-2.88;-0.84]	[-1.56;0.02]	[-0.05;1.50]	[-0.46;1.14]
Change in share legally employed (%-points)	-0.77	-0.21	0.74	2.78
	[-1.58;0.07]	[-0.98;0.59]	[-0.02;1.52]	[1.58;3.65]
Change in share incarcerated (%-points)	3.63	1.91	-0.70	-1.41
	[2.08;4.80]	[0.93;2.91]	[-1.59;-0.16]	[-2.34;0.37]

Notes:

1. In each row, I simulate alternative policies that achieve a one percentage point decrease in the monthly crime rate in Philadelphia. The crime rate is defined as the sum of the share of individuals participating in crime and employment/crime. In column (1), I simulate an extension on the average sentence length, which is achieved through a reduction in the release rate. In column (2), I simulate an increase in the incarceration rate through an increase in the number of police. In column (3), I simulate a reduction in the arrival rate of crime opportunities, through an increase in the number of police. For the simulations in columns (2) and (3), the appropriate changes in the incarceration rates and the arrival rates of crime opportunities are calculated within the model to match the elasticity of crime rate with respect to the number of police reported in Levitt (2004). In the last column, I simulate the introduction of a wage subsidy paid to legal workers on a monthly basis.

2. In the first row, I show the necessary change in the corresponding parameter (e.g. release rate). In the second row, I show the annual net cost of the policy per individual. The net cost of a policy includes the direct cost of the policy, minus the reduction in benefits paid to non-employed individuals, minus the increase in income taxes collected from legal workers, plus the reduction in prison expenditures, plus the change in the total cost of crime. All figures are in 2000 US dollars. The last three rows show the changes in the predicted outcomes relative to the baseline predicted outcomes of the model at the steady state using the parameter estimates from Table 11.

3. The direct cost per individual accounts for the specific cost of the policy. For the wage subsidy, the direct cost is equal to the total amount paid to workers on behalf of the wage subsidies. The direct cost of increasing the incarceration rate and reducing the arrival rate of crime opportunities is the cost of the extra police officers.

4. The numbers in brackets are the 95% bootstrap confidence intervals using 100 replications.

Lastly, I simulate an increase in legal earnings by introducing a wage subsidy. Besides their use for boosting legal employment (Phelps, 1994; Katz, 1996), wage subsidies have been proposed in the literature as a means of reducing criminal activity (Engelhardt, Rocheteau, and Rupert, 2008; Lochner, 2004). In order to achieve the desired reduction in crime, the government should pay a wage subsidy of $619 per month to legal workers in Philadelphia. As it is shown in column 5, the wage subsidy yields a large increase in legal employment (2.8 percentage points or 10.0%) and it does not generate an increase in the incarcerated population. Also, the share of non-employed individuals stays relatively flat. The net annual cost of this policy is $1,327 per individual in the model. Thus, the wage subsidy is more costly than policies targeting the criminal sector, mainly because of the high search frictions that characterize the legal labour sector in Philadelphia.

All things considered, I find that both the criminal sector and the legal labour sector provide tools to policy makers interested in reducing youth crime. Each policy has different implications on legal employment outcomes, non-employment, and the incarcerated population, which is ultimately reflected in the cost of each policy. Notably, usual interventions in the criminal sector, like increasing the apprehension rate, have sizable effects on crime; however, this is mainly achieved through a large increase in the population incarcerated. Instead, policies targeting the legal labour sector can reduce crime and generate a boost in legal employment without generating increases in the incarcerated population.

4.5 Conclusions

The school to work transition and the early career path can be difficult for disadvantaged youth. Their labour market is usually studied ignoring the presence of the criminal sector, and yet a large share of them participate in crime. In this chapter, I characterize the criminal and legal labour sectors for disadvantaged youth, and show that there are important interactions between youth crime and youth legal employment. I do so in the context of a two-sector search model, which allows for a rich set of interactions across sectors. On the one hand, individuals with higher earnings in the legal labour sector are less likely to participate in income crime due to the higher opportunity cost of being caught and sent to jail. On the other hand, individuals participating in income crime have less time to devote to either leisure or employment in the legal labour sector. They also have different legal job arrival and destruction rates, which are possibly additional sources of interaction between the two sectors. The model also accounts for simultaneous participation in the two sectors.

I estimate the model using a unique panel dataset on young serious offenders. I find that search frictions are quite important for this disadvantaged population. In particular, it takes them more than a year to receive a legal job offer when they are not participating in the criminal or legal labour sector. Legal jobs are characterized by short average durations, low earnings, and high destruction rates. The criminal sector is an attractive alternative for these individuals. I find that it offers an earnings premium relative to the legal labour sector but, not surprisingly, there is a high probability of incarceration and fewer opportunities in the legal labour sector. I find that there are sizable interactions across sectors, and that policies in one sector can have effects on the other sector. I provide evidence that a wage subsidy that targets the legal labour sector can reduce crime and boost legal employment among disadvantaged youth. A policy that reduces the arrival rate of crime opportunities, compared to policies usually discussed in the literature like extending the average sentence length, has the advantage of reducing crime without generating large increases in the incarcerated population.

Lastly, it is important to stress that I study youth who have already committed somewhat serious criminal offenses. 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 my results do not necessarily generalize to the population at large. In this sense, the factors that cause these young offenders to reduce crime may not be the same as those that prevent people from committing their first crime. Additionally, what helps to boost legal employment among young offenders, may not be as useful for increasing legal employment for individuals with no criminal history.

Bibliography

Agan, Amanda and Sonja Starr. 2017. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *Quarterly Journal of Economics* , forthcoming.

Albrecht, James, Lucas Navarro, and Susan Vroman. 2009. "The Effects of Labour Market Policies in an Economy with an Informal Sector." *The Economic Journal* 119 (539):1105–1129.

Albrecht, James, Monica Robayo-Abril, and Susan Vroman. 2017. "Public-Sector Employment in an Equilibrium Search and Matching Model." *The Economic Journal* , forthcoming.

Albrecht, James and Susan Vroman. 1992. "Dual Labor Markets, Efficiency Wages, and Search." *Journal of Labor Economics* 10 (4):438–461.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." In *The Economic Dimensions of Crime*. Springer, 13–68.

Bowlus, Audra J., Nicholas M. Kiefer, and George R. Neumann. 2001. "Equilibrium Search Models and the Transition from School to Work." *International Economic Review* 42 (2):317–343.

Braga, Anthony A. 2001. "The Effects of Hot Spots Policing on Crime." *Annals of the American Academy of Political and Social Science* 578 (1):104–125.

———. 2005. "Hot Spots Policing and Crime Prevention: A Systematic Review of Randomized Controlled Trials." *Journal of Experimental Criminology* 1 (3):317–342.

Braga, Anthony A., Andrew V. Papachristos, and David M. Hureau. 2014. "The Effects of Hot Spots Policing on Crime: An Updated Systematic Review and Meta-Analysis." *Justice quarterly* 31 (4):633–663.

Braun, Christine. 2018. "Crime and the Minimum Wage."

Burdett, Kenneth, Ricardo Lagos, and Randall Wright. 2003. "Crime, Inequality, and Unemployment." *The American Economic Review* 93 (5):1764–1777.

———. 2004. "An On-the-Job Search Model of Crime, Inequality, and Unemployment." *International Economic Review* 45 (3):681–706.

Bushway, Shawn D. and Gary Sweeten. 2007. “Abolish Lifetime Bans for Ex-Felons.” *Criminology & Public Policy* 6 (4):697–706.

Cameron, Samuel. 1988. “The Economics of Crime Deterrence: A Survey of Theory and Evidence.” *Kyklos* 41 (2):301–323.

Cano-Urbina, Javier. 2015. “The Role of the Informal Sector in the Early Careers of Less-Educated Workers.” *Journal of Development Economics* 112:33–55.

Chang, Juin-Jen, Huei-Chung Lu, and Ping Wang. 2013. “Search for a Theory of Organized Crimes.” *European Economic Review* 62:130–153.

Dey, Matthew and Christopher Flinn. 2008. “Household Search and Health Insurance Coverage.” *Journal of Econometrics* 145 (1):43–63.

Durlauf, Steven N. and Daniel S. Nagin. 2010. “The Deterrent Effect of Imprisonment.” In *Controlling Crime: Strategies and Tradeoffs*. University of Chicago Press, 43–94.

Eckstein, Zvi and Gerard J. Van den Berg. 2007. “Empirical Labor Search: A Survey.” *Journal of Econometrics* 136 (2):531–564.

Eckstein, Zvi and Kenneth I. Wolpin. 1995. “Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model.” *The Review of Economic Studies* 62 (2):263–286.

Ehrlich, Isaac. 1973. “Participation in Illegitimate Activities: A Theoretical and Empirical Investigation.” *Journal of Political Economy* 81 (3):521–565.

Engelhardt, Bryan. 2010. “The Effect of Employment Frictions on Crime.” *Journal of Labor Economics* 28 (3):677–718.

Engelhardt, Bryan, Guillaume Rocheteau, and Peter Rupert. 2008. “Crime and the Labor Market: A Search Model with Optimal Contracts.” *Journal of Public Economics* 92 (10):1876–1891.

Flabbi, Luca and James Mabli. 2018. “Household Search or Individual Search: Does it Matter?” *Journal of Labor Economics* , forthcoming.

Flinn, Christopher and James J. Heckman. 1982. “New Methods for Analyzing Structural Models of Labor Force Dynamics.” *Journal of Econometrics* 18 (1):115–68.

Freeman, Richard B. 2000. “Disadvantaged Young Men and Crime.” In *Youth Employment and Joblessness in Advanced Countries*. University of Chicago Press, 215–246.

Freeman, Richard B. and David A. Wise. 1982. “The Youth Labor Market Problem: Its Nature Causes and Consequences.” In *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*. University of Chicago Press, 1–16.

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

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

Gourieroux, Christian, Alain Monfort, and Eric Renault. 1993. "Indirect Inference." *Journal of Applied Econometrics* 8 (S1):S85–S118.

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

Guler, Bulent, Fatih Guvenen, and Giovanni L. Violante. 2012. "Joint-Search Theory: New Opportunities and New Frictions." *Journal of Monetary Economics* 59 (4):352–369.

Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics*, vol. 3A, edited by Orley Ashenfelter and David Card, chap. 31. New York: North-Holland, 1865–2097.

Heller, Sara B. 2014. "Summer Jobs Reduce Violence among Disadvantaged Youth." *Science* 346 (6214):1219–1223.

Huang, Chien-Chieh, Derek Laing, and Ping Wang. 2004. "Crime and Poverty: A Search-theoretic Approach." *International Economic Review* 45 (3):909–938.

Hughes, Kristen A. 2006. "Justice Expenditure and Employment in the United States, 2003." *Bureau of Justice Statistics, NCJ212260. US Department of Justice, Washington, DC* .

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

İmrohoroğlu, Ayse, Antonio Merlo, and Peter Rupert. 2004. "What Accounts for the Decline in Crime?" *International Economic Review* 45 (3):707–729.

Katz, Lawrence F. 1996. "Wage Subsidies for the Disadvantaged." Tech. rep., National Bureau of Economic Research.

Kessler, Daniel and Steven D. Levitt. 1999. "Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation." *The Journal of Law and Economics* 42 (S1):343–364.

Kiefer, Nicholas M. and George R. Neumann. 1993. "Wage Dispersion with Homogeneity: The Empirical Equilibrium Search Model." *Contributions to Economic Analysis* 222:57–57.

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

Lain, Jonathan. 2016. "Discrimination in a Search-Match Model with Self-Employment." Working paper, Centre for the Study of African Economies, University of Oxford.

Levitt, Steven D. 2004. "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not." *The Journal of Economic Perspectives* 18 (1):163–190.

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

Machin, Stephen and Costas Meghir. 2004. "Crime and Economic Incentives." *Journal of Human Resources* 39 (4):958–979.

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

Miller, Ted R., Mark A. Cohen, and Brian Wiersema. 1996. *Victim Costs and Consequences: A New Look*. Washington, D.C.: U.S. Dept. of Justice, Office of Justice Programs, National Institute of Justice.

Mocan, H. Naci, Stephen C. Billups, and Jody Overland. 2005. "A Dynamic Model of Differential Human Capital and Criminal Activity." *Economica* 72 (288):655–681.

Mortensen, Dale T. 1986. "Job Search and Labor Market Analysis." In *Handbook of Labor Economics*, *Handbooks in Economics*, vol. 2, edited by Orley Ashenfelter and Richard Layard. New York: Elsevier Science, 849–919.

Mortensen, Dale T. and Christopher A. Pissarides. 1999. "New Developments in Models of Search in the Labor Market." Elsevier, 2567–2627.

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

Nagin, Daniel S. 1998. "Deterrence and Incapacitation." *The Handbook of Crime and Punishment* :345–368.

Narita, Renata. 2012. "Self Employment in Developing Countries: a Search-Equilibrium Approach." Working paper, World Bank.

Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108 (5):937–975.

Phelps, Edmund S. 1994. "Low-Wage Employment Subsidies versus the Welfare State." *The American Economic Review* 84 (2):54–58.

Rajkumar, Andrew S. and Michael T. French. 1997. "Drug Abuse, Crime Costs, and the Economic Benefits of Treatment." *Journal of Quantitative Criminology* 13 (3):291–323.

Raphael, Steven and Melissa Sills. 2007. "Urban Crime, Race, and the Criminal Justice System in the United States." *A Companion to Urban Economics* :515–535.

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

Schmitt, John and Kris Warner. 2011. "Ex-Offenders and the Labor Market." *WorkingUSA* 14 (1):87–109.

Schnepel, Kevin T. 2016. "Good Jobs and Recidivism." *The Economic Journal* , forthcoming.

Sickles, Robin C. and Jenny Williams. 2008. "Turning from Crime: A Dynamic Perspective." *Journal of Econometrics* 145 (1):158–173.

Stephan, James J. 2001. *State Prison Expenditures, 2001*. Washington, DC: U.S. Dept. of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Uggen, Christopher. 2000. "Work as a Turning Point in the Life Course of Criminals: A Duration Model of Age, Employment, and Recidivism." *American Sociological Review* 65 (4):529–546.

Uggen, Christopher, Mike Vuolo, Sarah Lageson, Ebony Ruhland, and Hilary K. Whitmam. 2014. "The Edge of Stigma: An Experimental Audit of the Effects of Low-Level Criminal Records on Employment." *Criminology* 52 (4):627–654.

Wilson, David B., Catherine A. Gallagher, Mark B. Coggeshall, and Doris L. MacKenzie. 1999. "A Quantitative Review and Description of Corrections-based Education, Vocation, and Work Programs." *Corrections Management Quarterly* 3:8–18.

Wolpin, Kenneth I. 1987. "Estimating a Structural Search Model: The Transition from School to Work." *Econometrica* 55 (4):801–817.

Appendix A

Chapter 2 Appendices

A.1 Factor Model for Skills

Identification of the measurement/skills model of equations (2.4) and (2.5) follows from the analysis in Carneiro, Hansen, and Heckman (2003) and Cooley Fruehwirth, Navarro, and Takahashi (2016). The argument roughly follows from first (conditionally) demeaning the measurements, which recovers the $\beta' s$. The loadings (i.e., the δ's) are then identified by taking covariances between different cognitive measures and between different social/emotional measures. The marginal distributions of θ_i^{cog} and $\left\{\xi_{j,i,t_i}^{cog}\right\}_{j=1}^{J}$, as well as those of θ_i^{emo} and $\left\{\left\{\xi_{k,i,t}^{emo}\right\}_{t=t_i}^{T_i}\right\}_{k=1}^{K}$ are non-parametrically identified from a theorem of Kotlarski (1967) using deconvolution arguments. The correlation between θ_i^{cog} and θ_i^{emo} follows directly from the covariance between cognitive and social/emotional measures.

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

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

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

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

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

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

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

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

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

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

A.2 Additional Results

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

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

Continued on next page...

Table A.1 – continued from previous page.

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

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

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

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

Continued on next page...

Table A.2 – continued from previous page.

Variable	Excluding Drug Use		Enrollment Based on Attendance	
	(1)		(2)	
	Educ.	Crime	Educ.	Crime
		(0.052)		
Lagged Crime		0.194		0.156
		(0.013)		*(0.013)*
Cognitive Factor	0.037	0.058	-0.001	0.018
	(0.023)	*(0.025)*	*(0.018)*	*(0.024)*
Social/Emotional Factor	0.007	-0.122	0.009	-0.084
	(0.014)	*(0.015)*	*(0.011)*	*(0.015)*
Years of Education (alternative)			0.016	0.004
			(0.004)	*(0.004)*
Lagged Enrollment (alternative)			0.084	
			(0.010)	
Enrolment (alternative)				0.098
				(0.056)
Rho	-0.115		-0.228	
	(0.104)		*(0.124)*	
Observations	5,190	5,190	5,097	5,097

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In column (1) we do not include drug use as an independent regressor. In column (2), enrollment is redefined as attending school for at least nine months in a year.

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

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

Continued on next page...

Table A.3 – continued from previous page.

Variable	Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only	
	(1)		(2)		(3)		(4)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Enrollment				0.087		0.081		0.112
				(0.049)		*(0.049)*		*(0.055)*
Lagged Crime				0.156		0.158		0.131
				(0.013)		*(0.013)*		*(0.014)*
Cognitive Factor	0.036	0.016	0.035	0.012	0.036	0.014	0.038	0.013
	(0.023)	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*
Social/Emotional Factor	0.006	-0.080	0.007	-0.081	0.006	-0.080	0.009	-0.082
	(0.014)	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*
Years of Crime Age1	-0.006	0.022						
	(0.004)	*(0.004)*						
Years of Crime Age2	-0.007	0.020						
	(0.003)	*(0.003)*						
Years of Education Age1	0.002	-0.001						
	(0.005)	*(0.005)*						
Years of Education Age2	0.009	-0.002						
	(0.004)	*(0.005)*						
Lagged Enrollment Age1	0.229							
	(0.020)							
Lagged Enrollment Age2	0.170							
	(0.016)							
Enrollment Age1		0.065						
		(0.050)						
Enrollment Age2		0.046						
		(0.057)						
Lagged Crime Age1		0.156						
		(0.019)						
Lagged Crime Age2		0.160						
		(0.017)						
Years of Crime Squared			0.001	0.001				
			(0.001)	*(0.001)*				
Years of Crime: 0 to 4					-0.017	0.023		
					(0.008)	*(0.008)*		
Years of Crime: 5 to 9					-0.012	0.021		
					(0.004)	*(0.004)*		
Years of Crime: 10 or more					-0.008	0.021		
					(0.003)	*(0.003)*		
Years of Crime Age of Entry 14							-0.051	0.027
							(0.006)	*(0.007)*
Years of Crime Age of Entry 15							-0.026	0.043

Continued on next page...

Table A.3 – continued from previous page.

Variable	Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only	
	(1)		(2)		(3)		(4)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
							(0.005)	*(0.006)*
Years of Crime Age of Entry 16							-0.022	0.042
							(0.006)	*(0.006)*
Years of Crime Age of Entry 17							0.002	0.053
							(0.006)	*(0.007)*
Years of Crime Age of Entry 18							0.001	0.062
							(0.012)	*(0.012)*
Rho	-0.074		-0.142		-0.127		-0.195	
	(0.111)		*(0.107)*		*(0.106)*		*(0.122)*	
Observations	5,190	5,190	5,190	5,190	5,190	5,190	5,190	5,190

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

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

Variable	**Contemporaneous Effect of Crime on Education – Not Instrumenting**		**Contemporaneous Effect of Crime on Education**	
	(1)		(2)	
	Educ.	Crime	Educ.	Crime
Phoenix	0.043	0.051	0.041	-0.062
	(0.021)	*(0.019)*	*(0.021)*	*(0.062)*
Hispanic	-0.023	-0.023	-0.023	-0.023
	(0.015)	*(0.015)*	*(0.015)*	*(0.015)*
Black	0.027	-0.028	0.027	-0.028
	(0.017)	*(0.018)*	*(0.017)*	*(0.018)*
Other	0.035	-0.025	0.035	-0.028
	(0.027)	*(0.030)*	*(0.027)*	*(0.030)*
Female	0.067	-0.098	0.069	-0.097
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*
Non-intact Family	-0.053	0.026	-0.054	0.026
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*
Siblings	-0.002	0.003	-0.002	0.003
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*
Age	-0.074	-0.040	-0.072	-0.036
	(0.006)	*(0.004)*	*(0.006)*	*(0.004)*
Certainty of Punishment	0.005	-0.022	0.006	-0.021
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.018	0.006	-0.018	0.007
	(0.007)	*(0.007)*	*(0.007)*	*(0.007)*
Family Crime	-0.014	0.149	-0.018	0.148
	(0.018)	*(0.015)*	*(0.017)*	*(0.015)*
Drug Use	-0.029	0.225	-0.035	0.225
	(0.020)	*(0.010)*	*(0.019)*	*(0.010)*
Unemployment Rate	0.019	0.015	0.019	0.022
	(0.006)	*(0.005)*	*(0.006)*	*(0.006)*
Future Outlook Inventory	0.021	-0.022	0.022	-0.022
	(0.011)	*(0.011)*	*(0.011)*	*(0.011)*
Years of Crime	-0.010	0.020	-0.011	0.020
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*
Years of Education	0.007	-0.001	0.007	-0.000
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*
Schools per Young Person	0.301		0.296	
	(0.071)		*(0.071)*	
Lagged Enrollment	0.187		0.186	
	(0.012)		*(0.012)*	

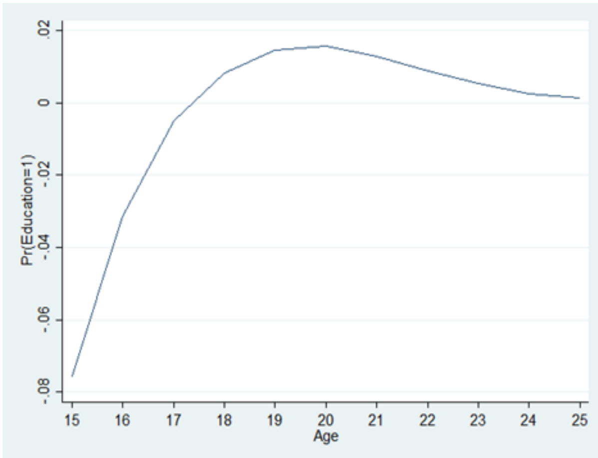
Continued on next page...

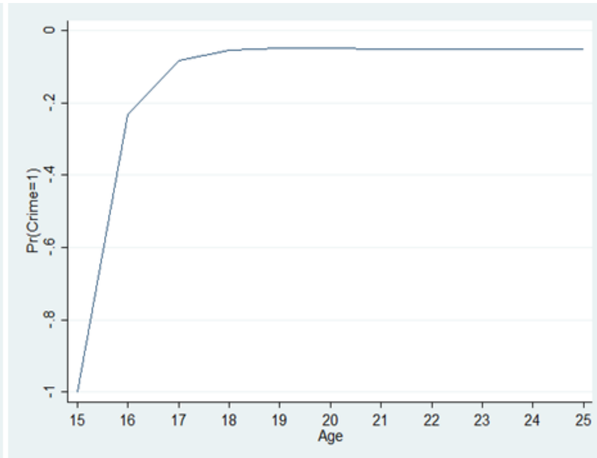
Table A.4 – continued from previous page.

Variable	**Contemporaneous Effect of Crime on Education – Not Instrumenting**		**Contemporaneous Effect of Crime on Education**	
	(1)		(2)	
	Educ.	Crime	Educ.	Crime
Lagged Crime		0.160		0.159
		(0.012)		*(0.012)*
Cognitive Factor	0.036	0.016	0.036	0.017
	(0.023)	*(0.024)*	*(0.023)*	*(0.024)*
Social/Emotional Factor	0.015	-0.079	0.017	-0.078
	(0.015)	*(0.015)*	*(0.015)*	*(0.015)*
Crime	0.097		0.120	
	(0.058)		*(0.056)*	
Lagged State Arrest Rate				-1.815
				(0.950)
Rho	-0.244		-0.186	
	(0.138)		*(0.141)*	
Observations	5,190	5,190	5,190	5,190

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In column (1) we change the direction of the contemporaneous effect; we estimate the contemporaneous effect of crime on education. In column (2) we add the lagged state arrest rate as an exclusion in the crime equation.

Figure A.1: No Crime at Age 15 - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect

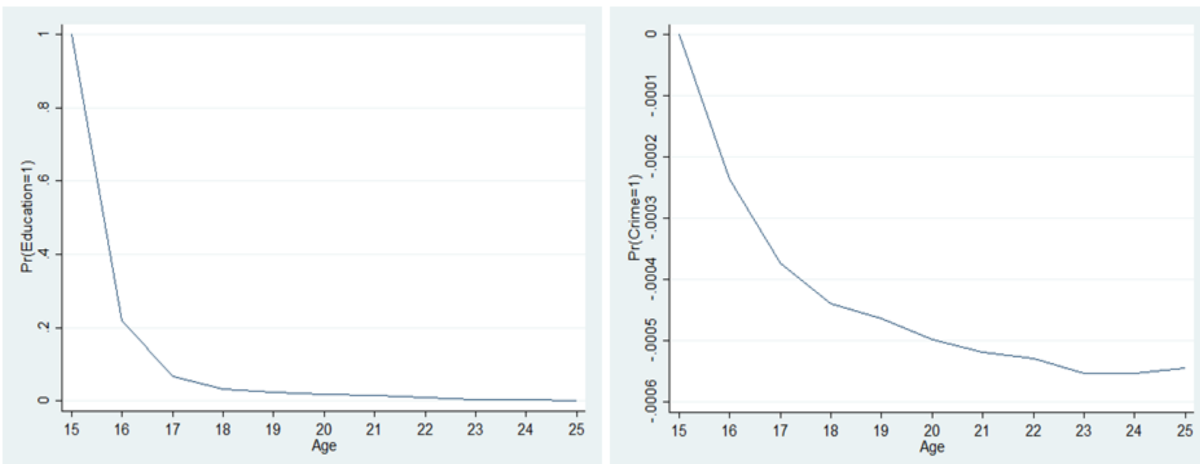




Notes:
1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are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

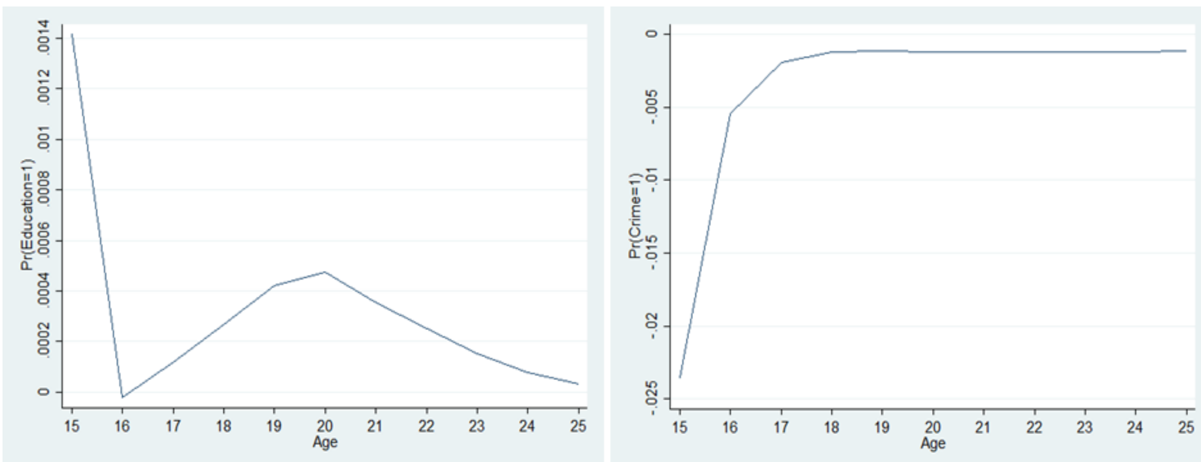
A.3 Results for Violent Crime, Drug-Related Crime, and Property Crime Samples

Figure A.2: Enrolled at Age 15 - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



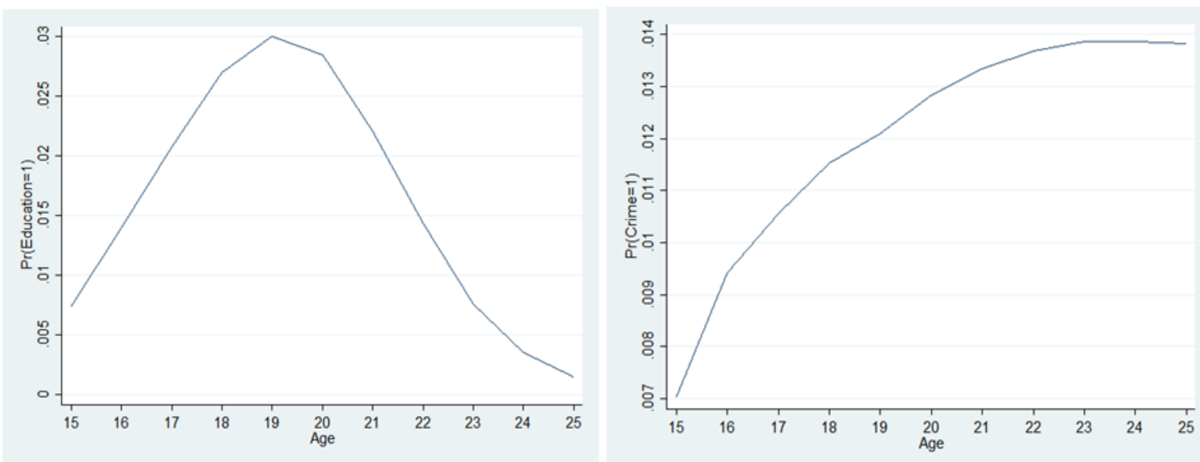
Notes:
1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are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

Figure A.3: Increase in Certainty of Punishment at Age 15 - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



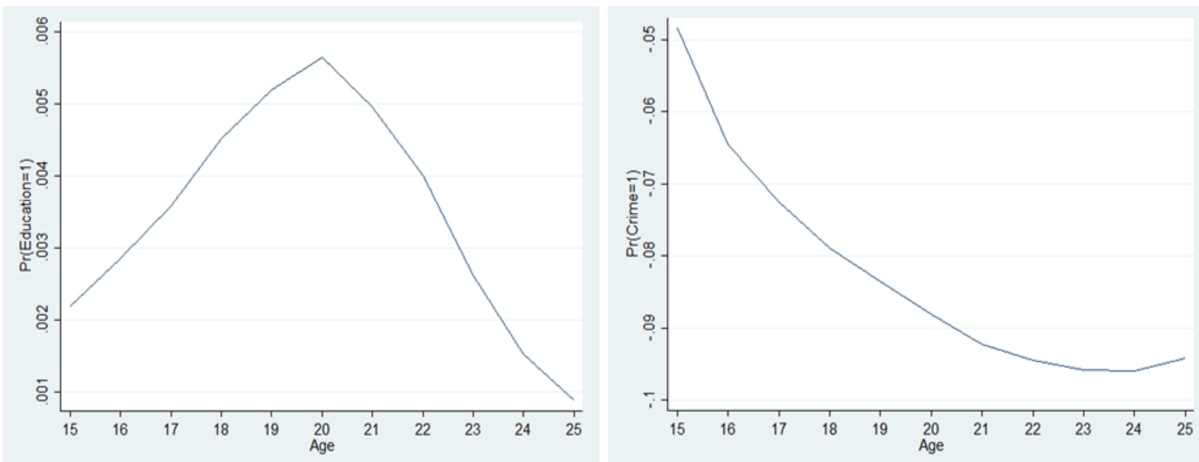
Notes:
1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are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

Figure A.4: Cognitive Factor 25th versus 75th Percentile - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



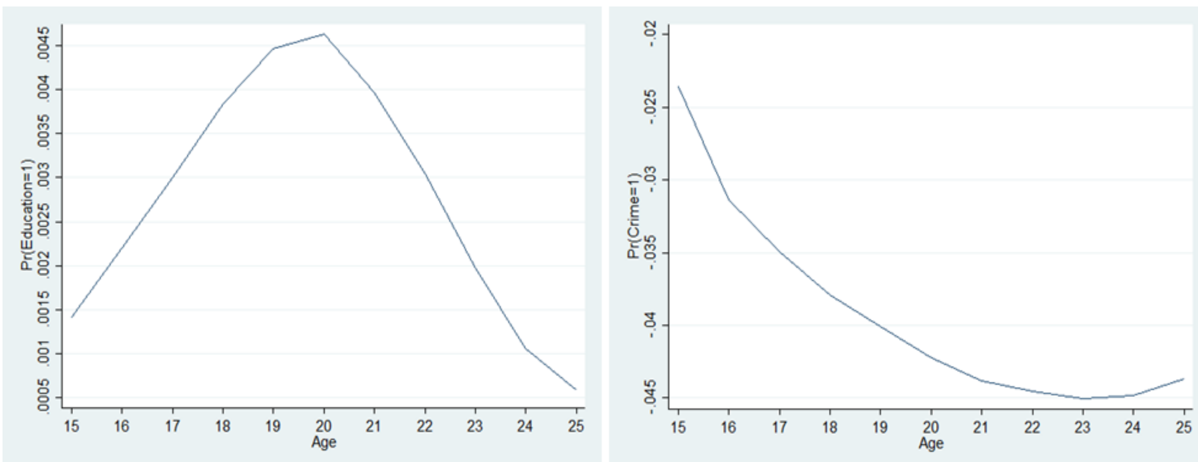
Notes:
1. The figures are based on the overall crime category.
2. For each simulation, the exogenous variables are set to their median level at age 15. We then draw pairs of errors for the crime and education equations from the estimated bivariate normal distribution. Crime and education decisions are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

Figure A.5: Social/Emotional Factor 25th versus 75th Percentile - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



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

Figure A.6: Increase in Certainty of Punishment (Permanent) - Effect on Average Probability of Education and Crime - Alternative Contemporaneous Effect



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

Table A.5: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime)

Variable	**Baseline**		**Controls**		**Uncorrelated Errors**		**No Dynamics**		**Not Instrumenting**		**Cognitive and Social/Emotional Skills**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.046	-0.014	0.035	-0.017	0.047	-0.005	0.047	-0.005	0.093	-0.012	0.033	-0.014
	(0.021)	*(0.015)*	*(0.019)*	*(0.014)*	*(0.021)*	*(0.015)*	*(0.022)*	*(0.016)*	*(0.018)*	*(0.015)*	*(0.022)*	*(0.016)*
Hispanic	-0.025	-0.022			-0.023	-0.024	-0.030	-0.028	-0.025	-0.022	-0.011	-0.027
	(0.015)	*(0.012)*			*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*
Black	0.026	-0.007			0.027	-0.004	0.045	-0.022	0.026	-0.006	0.043	-0.012
	(0.017)	*(0.014)*			*(0.017)*	*(0.014)*	*(0.018)*	*(0.015)*	*(0.017)*	*(0.014)*	*(0.018)*	*(0.014)*
Other	0.034	-0.015			0.036	-0.011	0.039	-0.022	0.036	-0.014	0.042	-0.010
	(0.027)	*(0.024)*			*(0.028)*	*(0.024)*	*(0.029)*	*(0.025)*	*(0.028)*	*(0.025)*	*(0.028)*	*(0.025)*
Female	0.060	-0.103	0.056	-0.099	0.059	-0.100	0.067	-0.142	0.060	-0.103	0.053	-0.099
	(0.014)	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*
Non-intact Family	-0.053	0.033			-0.053	0.029	-0.055	0.038	-0.054	0.032	-0.051	0.033
	(0.015)	*(0.013)*			*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Siblings	-0.002	-0.000			-0.002	-0.000	-0.003	0.001	-0.002	-0.000	-0.002	0.000
	(0.002)	*(0.002)*			*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.081	-0.000	-0.084	-0.009	-0.081	-0.010	-0.104	0.009	-0.083	-0.002	-0.080	0.001
	(0.004)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.003)*	*(0.002)*	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.005)*
Certainty of Punishment	0.003	-0.010			0.003	-0.010	0.005	-0.015	0.003	-0.010	0.003	-0.008
	(0.003)	*(0.002)*			*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Children	-0.016	0.009			-0.016	0.007	-0.033	0.019	-0.016	0.008	-0.016	0.007
	(0.008)	*(0.006)*			*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*
Family Crime	0.001	0.084			0.001	0.085	-0.005	0.107	0.003	0.084	-0.000	0.082
	(0.015)	*(0.010)*			*(0.015)*	*(0.010)*	*(0.015)*	*(0.011)*	*(0.015)*	*(0.010)*	*(0.015)*	*(0.010)*
Drug Use	0.006	0.213			0.005	0.214	-0.005	0.255	0.007	0.213	0.006	0.203
	(0.012)	*(0.010)*			*(0.012)*	*(0.010)*	*(0.012)*	*(0.009)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.010)*
Unemployment Rate	0.020	0.002	0.021	0.002	0.021	0.006	0.022	0.001	0.037	0.003	0.021	0.001
	(0.006)	*(0.004)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.005)*	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Future Outlook Inventory	0.018	-0.008			0.019	-0.006	0.022	-0.008	0.016	-0.008	0.024	0.007
	(0.011)	*(0.009)*			*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.012)*	*(0.010)*
Years of Crime	-0.011	0.022	-0.012	0.040	-0.011	0.021			-0.011	0.022	-0.011	0.020
	(0.004)	*(0.003)*	*(0.003)*	*(0.003)*	*(0.004)*	*(0.003)*			*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Years of Education	0.007	-0.008	0.011	-0.011	0.007	-0.006			0.006	-0.008	0.006	-0.009
	(0.004)	*(0.003)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.003)*			*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Cognitive Factor	0.049	0.022			0.048	0.026	0.061	0.034	0.048	0.023		
	(0.023)	*(0.018)*			*(0.023)*	*(0.018)*	*(0.024)*	*(0.019)*	*(0.023)*	*(0.018)*		
Social/Emotional Factor	0.006	-0.028			0.006	-0.028	0.014	-0.049	0.006	-0.028		
	(0.014)	*(0.011)*			*(0.014)*	*(0.011)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.011)*		

Continued on next page...

Table A.5 – continued from previous page.

Variable	**Baseline**		**Controls**		**Uncorrelated Errors**		**No Dynamics**		**Not Instrumenting**		**Cognitive and Social/Emotional Skills**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Schools per Young Person	0.326		0.324		0.312		0.323				0.312	
	(0.072)		*(0.072)*		*(0.072)*		*(0.073)*				*(0.071)*	
Lagged Enrollment	0.191		0.194		0.192				0.192		0.186	
	(0.012)		*(0.012)*		*(0.012)*				*(0.013)*		*(0.013)*	
Enrollment		0.077		0.063		0.000		0.154		0.062		0.088
		(0.036)		*(0.038)*		*(0.011)*		*(0.039)*		*(0.038)*		*(0.034)*
Lagged Crime		0.097		0.149		0.098				0.098		0.092
		(0.010)		*(0.011)*		*(0.010)*				*(0.010)*		*(0.010)*
WASI Reasoning Score											-0.001	0.001
											(0.006)	*(0.005)*
WASI Vocabulary Score											0.002	-0.002
											(0.007)	*(0.006)*
Stroop: Color/Word											0.005	-0.007
											(0.007)	*(0.005)*
Stroop: Word											0.010	-0.004
											(0.007)	*(0.006)*
Stroop: Color											-0.004	0.016
											(0.008)	*(0.006)*
Trail-Making: Part B											-0.018	0.004
											(0.007)	*(0.005)*
Trail-Making: Part A											-0.004	0.004
											(0.007)	*(0.005)*
WAI - Impulse Response											-0.008	-0.008
											(0.008)	*(0.006)*
WAI - Suppression of Aggression											0.011	-0.025
											(0.007)	*(0.006)*
WAI - Consideration of Others											0.000	-0.012
											(0.006)	*(0.005)*
PSMI - Self Reliance											-0.013	0.020
											(0.010)	*(0.008)*
PSMI - Identity											0.036	-0.029
											(0.010)	*(0.008)*
PSMI - Work Orientation											-0.026	0.010
											(0.009)	*(0.007)*
Rho	-0.289		-0.198				-0.548		-0.230		-0.320	
	(0.130)		(0.109)				(0.140)		(0.137)		*(0.129)*	
Observations	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074

Continued on next page...

Table A.5 – continued from previous page.

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime

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

Table A.6: Average Marginal Effects from Probits for Crime and Education (Violent Crime)

Variable	**Baseline**		**Controls**		**Uncorrelated Errors**		**No Dynamics**		**Not Instrumenting**		**Cognitive and Social/Emotional Skills**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.044	0.037	0.032	0.036	0.045	0.045	0.049	0.032	0.092	0.040	0.035	0.043
	(0.021)	*(0.020)*	*(0.019)*	*(0.018)*	*(0.021)*	*(0.019)*	*(0.021)*	*(0.021)*	*(0.018)*	*(0.020)*	*(0.022)*	*(0.021)*
Hispanic	-0.023	-0.018			-0.023	-0.020	-0.027	-0.021	-0.024	-0.019	-0.010	-0.026
	(0.015)	*(0.015)*			*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*
Black	0.027	-0.032			0.027	-0.030	0.043	-0.040	0.027	-0.031	0.043	-0.032
	(0.017)	*(0.018)*			*(0.017)*	*(0.018)*	*(0.018)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.018)*	*(0.019)*
Other	0.038	-0.015			0.037	-0.013	0.040	-0.011	0.039	-0.014	0.046	-0.007
	(0.027)	*(0.030)*			*(0.027)*	*(0.030)*	*(0.028)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*
Female	0.056	-0.079	0.052	-0.071	0.056	-0.076	0.070	-0.148	0.055	-0.078	0.051	-0.075
	(0.015)	*(0.017)*	*(0.014)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*
Non-intact Family	-0.048	0.014			-0.048	0.010	-0.051	0.021	-0.049	0.013	-0.046	0.015
	(0.015)	*(0.016)*			*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*
Siblings	-0.002	0.004			-0.002	0.004	-0.003	0.006	-0.002	0.004	-0.002	0.004
	(0.002)	*(0.003)*			*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*
Age	-0.080	-0.030	-0.084	-0.032	-0.080	-0.039	-0.104	-0.019	-0.082	-0.032	-0.079	-0.029
	(0.004)	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.004)*	*(0.002)*	*(0.009)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.007)*
Certainty of Punishment	0.003	-0.019			0.003	-0.019	0.005	-0.024	0.003	-0.019	0.002	-0.016
	(0.003)	*(0.003)*			*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.017	0.008			-0.017	0.007	-0.032	0.010	-0.017	0.008	-0.017	0.003
	(0.007)	*(0.008)*			*(0.007)*	*(0.008)*	*(0.008)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.007)*
Family Crime	0.005	0.130			0.005	0.132	-0.001	0.154	0.006	0.130	0.004	0.126
	(0.015)	*(0.015)*			*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.014)*
Drug Use	-0.004	0.159			-0.004	0.159	-0.013	0.190	-0.004	0.159	-0.004	0.138
	(0.011)	*(0.011)*			*(0.011)*	*(0.011)*	*(0.012)*	*(0.012)*	*(0.011)*	*(0.011)*	*(0.012)*	*(0.011)*
Unemployment Rate	0.021	0.005	0.021	0.004	0.021	0.009	0.022	0.004	0.037	0.006	0.021	0.005
	(0.006)	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.006)*
Future Outlook Inventory	0.018	-0.024			0.017	-0.024	0.025	-0.028	0.016	-0.024	0.022	0.012
	(0.011)	*(0.012)*			*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.012)*	*(0.013)*
Years of Crime	-0.007	0.021	-0.008	0.033	-0.007	0.021			-0.007	0.021	-0.007	0.017
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*			*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*
Years of Education	0.007	0.001	0.012	-0.010	0.006	0.003			0.006	0.002	0.006	-0.000
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*			*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Cognitive Factor	0.043	0.030			0.045	0.034	0.048	0.054	0.042	0.032		
	(0.023)	*(0.024)*			*(0.023)*	*(0.024)*	*(0.024)*	*(0.024)*	*(0.023)*	*(0.024)*		
Social/Emotional Factor	0.007	-0.074			0.008	-0.074	0.018	-0.111	0.008	-0.074		
	(0.014)	*(0.015)*			*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.015)*		

Continued on next page...

Table A.6 – continued from previous page.

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Schools per Young Person	0.332		0.334		0.322		0.326				0.319	
	(0.071)		*(0.071)*		*(0.071)*		*(0.073)*				*(0.071)*	
Lagged Enrollment	0.189		0.191		0.190				0.190		0.185	
	(0.012)		*(0.012)*		*(0.012)*				*(0.012)*		*(0.012)*	
Enrollment		0.104		0.101		0.033		0.199		0.083		0.106
		(0.048)		*(0.051)*		*(0.014)*		*(0.061)*		*(0.051)*		*(0.047)*
Lagged Crime		0.142		0.188		0.144				0.143		0.125
		(0.012)		*(0.013)*		*(0.012)*				*(0.012)*		*(0.012)*
WASI Reasoning Score											0.000	-0.005
											(0.006)	*(0.007)*
WASI Vocabulary Score											-0.002	-0.009
											(0.007)	*(0.007)*
Stroop: Color/Word											0.004	-0.002
											(0.007)	*(0.007)*
Stroop: Word											0.010	-0.003
											(0.007)	*(0.008)*
Stroop: Color											-0.002	0.008
											(0.008)	*(0.008)*
Trail-Making: Part B											-0.016	-0.007
											(0.007)	*(0.007)*
Trail-Making: Part A											-0.003	-0.001
											(0.007)	*(0.007)*
WAI - Impulse Response											-0.009	-0.025
											(0.007)	*(0.008)*
WAI - Suppression of Aggression											0.011	-0.059
											(0.007)	*(0.007)*
WAI - Consideration of Others											0.001	-0.027
											(0.006)	*(0.006)*
PSMI - Self Reliance											-0.014	0.013
											(0.010)	*(0.011)*
PSMI - Identity											0.035	-0.008
											(0.010)	*(0.011)*
PSMI - Work Orientation											-0.022	0.004
											(0.009)	*(0.009)*
Rho	-0.156		-0.157				-0.348		-0.109		-0.159	
	(0.104)		(0.102)				(0.137)		(0.108)		*(0.104)*	
Observations	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232

Continued on next page...

Table A.6 – continued from previous page.

Variable	**Baseline**		**Controls**		**Uncorrelated Errors**		**No Dynamics**		**Not Instrumenting**		**Cognitive and Social/Emotional Skills**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime

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

Table A.7: Average Marginal Effects from Probits for Crime and Education (Property Crime)

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

Continued on next page...

Table A.7 – continued from previous page.

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Schools per Young Person	0.327		0.328		0.329		0.310				0.314	
	(0.071)		*(0.072)*		*(0.071)*		*(0.074)*				*(0.071)*	
Lagged Enrollment	0.190		0.192		0.191				0.191		0.186	
	(0.012)		*(0.012)*		*(0.012)*				*(0.012)*		*(0.012)*	
Enrollment		-0.022		-0.019		-0.003		-0.017		-0.035		-0.011
		(0.040)		*(0.041)*		*(0.012)*		*(0.077)*		*(0.042)*		*(0.040)*
Lagged Crime		0.144		0.196		0.144				0.144		0.132
		(0.010)		*(0.011)*		*(0.010)*				*(0.010)*		*(0.010)*
WASI Reasoning Score											0.001	0.001
											(0.006)	*(0.006)*
WASI Vocabulary Score											-0.003	0.013
											(0.007)	*(0.006)*
Stroop: Color/Word											0.003	-0.011
											(0.007)	*(0.006)*
Stroop: Word											0.008	-0.002
											(0.007)	*(0.006)*
Stroop: Color											-0.001	0.005
											(0.008)	*(0.007)*
Trail-Making: Part B											-0.016	-0.007
											(0.007)	*(0.006)*
Trail-Making: Part A											-0.004	0.000
											(0.007)	*(0.006)*
WAI - Impulse Response											-0.007	-0.030
											(0.007)	*(0.006)*
WAI - Suppression of Aggression											0.013	-0.029
											(0.007)	*(0.006)*
WAI - Consideration of Others											0.001	-0.022
											(0.006)	*(0.005)*
PSMI - Self Reliance											-0.013	0.012
											(0.010)	*(0.009)*
PSMI - Identity											0.035	-0.012
											(0.010)	*(0.009)*
PSMI - Work Orientation											-0.024	-0.015
											(0.009)	*(0.008)*
Rho	0.061		0.014				0.039		0.100		0.041	
	(0.118)		(0.107)				(0.202)		(0.123)		*(0.119)*	
Observations	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232

Continued on next page...

Table A.7 – continued from previous page.

Variable	Baseline		Controls		Uncorrelated Errors		No Dynamics		Not Instrumenting		Cognitive and Social/Emotional Skills	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime

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

Table A.8: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime) - Robustness Checks 1 and 2

Variable	**Excluding Drug Use**		**Choices While in Jail (1)**		**Choices While in Jail (2)**		**Choices While in Jail (3)**		**Enrollment Based on Attendance**	
	(1)		(2)		(3)		(4)		(5)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.046	-0.013	0.067	0.001	0.052	-0.005	0.051	-0.007	0.014	-0.005
	(0.021)	*(0.016)*	*(0.021)*	*(0.014)*	*(0.019)*	*(0.014)*	*(0.019)*	*(0.014)*	*(0.018)*	*(0.015)*
Hispanic	-0.024	-0.030	-0.032	-0.014	-0.005	-0.019	-0.005	-0.019	-0.020	-0.023
	(0.015)	*(0.012)*	*(0.015)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.012)*	*(0.012)*
Black	0.026	-0.020	-0.003	0.002	0.045	-0.006	0.043	-0.008	0.001	-0.007
	(0.017)	*(0.015)*	*(0.017)*	*(0.013)*	*(0.016)*	*(0.013)*	*(0.016)*	*(0.013)*	*(0.014)*	*(0.014)*
Other	0.033	-0.036	0.012	-0.007	0.033	-0.009	0.032	-0.009	-0.022	-0.012
	(0.027)	*(0.026)*	*(0.027)*	*(0.023)*	*(0.027)*	*(0.023)*	*(0.027)*	*(0.023)*	*(0.022)*	*(0.025)*
Female	0.060	-0.103	0.114	-0.092	0.023	-0.078	0.024	-0.076	0.013	-0.100
	(0.014)	*(0.015)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.011)*	*(0.014)*
Non-intact Family	-0.053	0.035	-0.062	0.021	-0.033	0.019	-0.033	0.019	-0.007	0.030
	(0.015)	*(0.014)*	*(0.015)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.012)*	*(0.013)*
Siblings	-0.002	-0.002	-0.002	-0.001	-0.002	-0.001	-0.002	-0.001	-0.004	-0.001
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.081	-0.006	-0.069	-0.010	-0.076	-0.004	-0.075	-0.005	-0.048	-0.013
	(0.004)	*(0.006)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.005)*	*(0.003)*	*(0.004)*
Certainty of Punishment	0.003	-0.013	0.007	-0.008	-0.000	-0.007	-0.000	-0.007	0.006	-0.010
	(0.003)	*(0.002)*	*(0.003)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Children	-0.017	0.004	-0.013	0.010	-0.015	0.013	-0.013	0.013	-0.024	0.012
	(0.008)	*(0.006)*	*(0.008)*	*(0.005)*	*(0.007)*	*(0.005)*	*(0.007)*	*(0.005)*	*(0.008)*	*(0.006)*
Family Crime	0.002	0.108	-0.029	0.078	0.005	0.069	0.004	0.069	0.009	0.083
	(0.015)	*(0.011)*	*(0.014)*	*(0.009)*	*(0.014)*	*(0.009)*	*(0.014)*	*(0.009)*	*(0.011)*	*(0.010)*
Drug Use			0.056	0.228	-0.038	0.235	-0.039	0.235	-0.044	0.217
			(0.011)	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.009)*	*(0.010)*
Unemployment Rate	0.020	0.002	0.020	0.007	0.022	0.004	0.022	0.004	0.018	0.006
	(0.006)	*(0.005)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.005)*	*(0.004)*
Future Outlook Inventory	0.018	-0.023	0.019	-0.009	0.015	-0.007	0.015	-0.008	0.017	-0.006
	(0.011)	*(0.010)*	*(0.011)*	*(0.008)*	*(0.010)*	*(0.008)*	*(0.010)*	*(0.008)*	*(0.009)*	*(0.009)*
Years of Crime	-0.010	0.033	-0.021	0.023	-0.002	0.021	-0.003	0.023	-0.003	0.021
	(0.004)	*(0.003)*	*(0.004)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.003)*	*(0.003)*
Years of Education	0.007	-0.006	0.022	-0.004	-0.007	-0.003	-0.002	-0.003		
	(0.004)	*(0.004)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.004)*		
Schools per Young Person	0.325		0.146		0.322		0.324		0.141	
	(0.072)		*(0.072)*		*(0.067)*		*(0.067)*		*(0.065)*	
Lagged Enrollment	0.191		0.179		0.199		0.223			
	(0.012)		*(0.013)*		*(0.011)*		*(0.013)*			
Enrollment		0.069		-0.002						

Continued on next page...

Table A.8 – continued from previous page.

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance	
	(1)		(2)		(3)		(4)		(5)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
		(0.038)		*(0.012)*						
Lagged Crime		0.130		0.088		0.085		0.098		0.098
		(0.011)		*(0.010)*		*(0.010)*		*(0.012)*		*(0.010)*
Cognitive Factor	0.050	0.060	0.042	0.016	0.007	0.025	0.005	0.024	0.004	0.030
	(0.023)	*(0.019)*	*(0.023)*	*(0.017)*	*(0.022)*	*(0.017)*	*(0.022)*	*(0.017)*	*(0.019)*	*(0.018)*
Social/Emotional Factor	0.005	-0.066	0.014	-0.026	-0.011	-0.025	-0.012	-0.025	0.012	-0.031
	(0.014)	*(0.012)*	*(0.013)*	*(0.011)*	*(0.013)*	*(0.010)*	*(0.013)*	*(0.010)*	*(0.011)*	*(0.011)*
Jail					0.104	0.063	0.367	0.068		
					(0.012)	*(0.010)*	*(0.076)*	*(0.065)*		
Enrollment (alternative)						0.042		0.038		
						(0.035)		*(0.035)*		
Years of Crime Jail							0.004	-0.006		
							(0.007)	*(0.006)*		
Years of Education Jail							-0.019	0.002		
							(0.007)	*(0.005)*		
Lagged Enrollment Jail							-0.088			
							(0.023)			
Lagged Crime Jail								-0.034		
								(0.020)		
Enrollment Jail								-0.001		
								(0.019)		
Years of Education (alternative)									0.017	-0.002
									(0.004)	*(0.003)*
Lagged Enrollment (alternative)									0.086	
									(0.010)	
Enrolment (alternative)										-0.005
										(0.048)
Rho	-0.214		0.030		-0.166		-0.152		-0.051	
	(0.117)		*(0.041)*		*(0.121)*		*(0.121)*		*(0.165)*	
Observations	5,074	5,074	6,042	6,042	6,042	6,042	6,042	6,042	4,987	4,987

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In column (1) we do not include drug use as an independent regressor. In column (2), enrollment is set to zero if an individual did not attend a community school. In column (3), we condition on whether the individual is interviewed in jail, and in column (4) we interact the jail dummy with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail. In column (5) enrollment is redefined as attending school for at least nine months.

Table A.9: Average Marginal Effects from Probits for Crime and Education (Drug-Related Crime) - Robustness Checks 3 and 4

Variable	**Age-Varying Coefficients**		**Years of Crime: Quadratic**		**Years of Crime: Piecewise-linear**		**Years of Crime: Observed Experience Only**		**Contemporaneous Effect of Crime on Education – No Instrument**		**Contemporaneous Effect of Crime on Education**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.048	-0.012	0.045	-0.012	0.045	-0.012	0.045	-0.011	0.045	-0.063	0.045	-0.003
	(0.021)	*(0.015)*	*(0.021)*	*(0.015)*	*(0.021)*	*(0.015)*	*(0.021)*	*(0.015)*	*(0.021)*	*(0.047)*	*(0.021)*	*(0.015)*
Hispanic	-0.024	-0.021	-0.024	-0.022	-0.024	-0.022	-0.024	-0.024	-0.019	-0.023	-0.019	-0.023
	(0.015)	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*	*(0.015)*	*(0.012)*
Black	0.025	-0.007	0.026	-0.006	0.025	-0.006	0.023	-0.011	0.026	-0.005	0.026	-0.004
	(0.017)	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*	*(0.017)*	*(0.014)*
Other	0.036	-0.014	0.034	-0.015	0.033	-0.014	0.027	-0.011	0.035	-0.015	0.035	-0.014
	(0.027)	*(0.024)*	*(0.028)*	*(0.025)*	*(0.028)*	*(0.024)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*
Female	0.062	-0.101	0.060	-0.104	0.060	-0.103	0.058	-0.108	0.070	-0.100	0.070	-0.100
	(0.014)	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*
Non-intact Family	-0.054	0.033	-0.053	0.033	-0.053	0.033	-0.051	0.034	-0.057	0.028	-0.057	0.028
	(0.015)	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Siblings	-0.002	-0.000	-0.002	-0.000	-0.002	-0.000	-0.002	0.001	-0.002	-0.000	-0.002	-0.000
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.088	-0.001	-0.081	-0.000	-0.081	-0.000	-0.084	-0.006	-0.078	-0.008	-0.078	-0.010
	(0.005)	*(0.006)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Certainty of Punishment	0.003	-0.010	0.003	-0.010	0.003	-0.010	0.003	-0.011	0.005	-0.010	0.005	-0.010
	(0.003)	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Children	-0.016	0.009	-0.016	0.009	-0.016	0.009	-0.019	0.007	-0.018	0.008	-0.018	0.007
	(0.008)	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*	*(0.008)*	*(0.006)*
Family Crime	0.002	0.083	0.001	0.084	0.000	0.085	0.001	0.083	-0.019	0.084	-0.018	0.084
	(0.015)	*(0.010)*	*(0.015)*	*(0.010)*	*(0.015)*	*(0.010)*	*(0.015)*	*(0.010)*	*(0.017)*	*(0.010)*	*(0.017)*	*(0.010)*
Drug Use	0.007	0.212	0.006	0.213	0.006	0.213	0.008	0.216	-0.028	0.214	-0.026	0.214
	(0.012)	*(0.010)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.010)*	*(0.012)*	*(0.010)*	*(0.018)*	*(0.010)*	*(0.018)*	*(0.010)*
Unemployment Rate	0.020	0.003	0.020	0.002	0.020	0.002	0.021	0.004	0.019	0.010	0.019	0.007
	(0.006)	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.004)*
Future Outlook Inventory	0.020	-0.008	0.018	-0.008	0.018	-0.008	0.021	-0.005	0.020	-0.007	0.020	-0.007
	(0.011)	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*
Years of Crime			-0.017	0.029					-0.016	0.021	-0.016	0.021
			(0.008)	*(0.006)*					*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Years of Education			0.007	-0.008	0.007	-0.008	0.010	-0.007	0.008	-0.007	0.008	-0.007
			(0.004)	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*
Schools per Young Person	0.332		0.328		0.327		0.297		0.304		0.307	
	(0.072)		*(0.072)*		*(0.072)*		*(0.072)*		*(0.072)*		*(0.071)*	
Lagged Enrollment			0.191		0.191		0.182		0.188		0.189	
			(0.012)		*(0.012)*		*(0.013)*		*(0.013)*		*(0.013)*	
Enrollment				0.077		0.075		0.057				
				(0.035)		*(0.036)*		*(0.040)*				
Lagged Crime				0.096		0.096		0.089		0.096		0.096
				(0.010)		*(0.010)*		*(0.011)*		*(0.010)*		*(0.010)*
Cognitive Factor	0.047	0.020	0.049	0.023	0.049	0.023	0.048	0.022	0.043	0.024	0.044	0.024
	(0.023)	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*	*(0.023)*	*(0.018)*
Social/Emotional Factor	0.004	-0.029	0.005	-0.028	0.006	-0.028	0.007	-0.025	0.009	-0.024	0.009	-0.025
	(0.014)	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.014)*	*(0.011)*
Years of Crime Age1	-0.016	0.027										
	(0.006)	*(0.005)*										
Years of Crime Age2	-0.008	0.019										
	(0.004)	*(0.004)*										

Continued on next page...

Table A.9 – continued from previous page.

Variable	Age-Varying Coefficients		Years of Crime: Quadratic		Years of Crime: Piecewise-linear		Years of Crime: Observed Experience Only		Contemporaneous Effect of Crime on Education – No Instrument		Contemporaneous Effect of Crime on Education	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Education Age1	0.004	-0.008										
	(0.005)	*(0.004)*										
Years of Education Age2	0.009	-0.008										
	(0.004)	*(0.003)*										
Lagged Enrollment Age1	0.228											
	(0.020)											
Lagged Enrollment Age2	0.172											
	(0.016)											
Enrollment Age1		0.080										
		(0.036)										
Enrollment Age2		0.089										
		(0.045)										
Lagged Crime Age1		0.072										
		(0.015)										
Lagged Crime Age2		0.123										
		(0.015)										
Years of Crime Squared			0.001	-0.001								
			(0.002)	*(0.001)*								
Years of Crime: 0 to 4					-0.013	0.024						
					(0.004)	*(0.004)*						
Years of Crime: 5 to 9					-0.006	0.019						
					(0.005)	*(0.003)*						
Years of Crime: 10 or more					-0.088	0.119						
					(32.524)	*(23.080)*						
Years of Crime Age of Entry 14							-0.063	0.019				
							(0.013)	*(0.009)*				
Years of Crime Age of Entry 15							-0.035	0.023				
							(0.011)	*(0.008)*				
Years of Crime Age of Entry 16							-0.023	0.029				
							(0.008)	*(0.006)*				
Years of Crime Age of Entry 17							-0.000	0.046				
							(0.009)	*(0.006)*				
Years of Crime Age of Entry 18							-0.003	0.029				
							(0.019)	*(0.011)*				
Crime									0.147		0.137	
									(0.058)		*(0.062)*	
Lagged State Arrest Rate										-0.962		
										(0.727)		
Rho	-0.311		-0.287		-0.281		-0.225		-0.353		-0.379	
	(0.142)		*(0.130)*		*(0.131)*		*(0.145)*		*(0.162)*		*(0.155)*	
Observations	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5,074	5074	5074	5074	5074

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In column (1), coefficients are allowed to vary by age. Age1 is a dummy for ages 14 to 19, and Age2 is a dummy for ages 20 and above. In column (2) we use a quadratic function in criminal experience. In column (3) we use a piecewise-linear function of criminal experience: 0 to 4, 5 to 9, more than 10. In column (4) we use the criminal experience observed in the sample only, interacted with age of entry dummies. In column (5) we change the direction of the contemporaneous effect; we estimate the contemporaneous effect of crime on education. In column (6) we add the lagged state arrest rate as an exclusion in the crime equation.

Table A.10: Average Marginal Effects from Probits for Crime and Education (Violent Crime) - Robustness Checks 1 and 2

Variable	**Excluding Drug Use**		**Choices While in Jail (1)**		**Choices While in Jail (2)**		**Choices While in Jail (3)**		**Enrollment Based on Attendance**	
	(1)		(2)		(3)		(4)		(5)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.044	0.038	0.065	0.041	0.054	0.036	0.052	0.035	0.012	0.041
	(0.021)	*(0.021)*	*(0.021)*	*(0.018)*	*(0.019)*	*(0.019)*	*(0.019)*	*(0.019)*	*(0.018)*	*(0.020)*
Hispanic	-0.023	-0.023	-0.033	-0.021	-0.005	-0.025	-0.005	-0.026	-0.018	-0.015
	(0.015)	*(0.016)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.012)*	*(0.016)*
Black	0.027	-0.039	0.003	-0.035	0.043	-0.041	0.040	-0.041	0.007	-0.027
	(0.017)	*(0.019)*	*(0.017)*	*(0.017)*	*(0.016)*	*(0.017)*	*(0.016)*	*(0.017)*	*(0.013)*	*(0.018)*
Other	0.038	-0.028	0.011	-0.005	0.034	-0.007	0.032	-0.007	-0.019	-0.010
	(0.027)	*(0.031)*	*(0.027)*	*(0.028)*	*(0.026)*	*(0.028)*	*(0.026)*	*(0.028)*	*(0.022)*	*(0.030)*
Female	0.056	-0.079	0.116	-0.078	0.017	-0.059	0.018	-0.057	0.002	-0.070
	(0.015)	*(0.017)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.017)*	*(0.012)*	*(0.017)*
Non-intact Family	-0.048	0.011	-0.055	0.010	-0.028	0.005	-0.029	0.004	-0.003	0.010
	(0.015)	*(0.016)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.012)*	*(0.016)*
Siblings	-0.002	0.002	-0.003	0.003	-0.002	0.002	-0.001	0.002	-0.005	0.004
	(0.002)	*(0.003)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.003)*
Age	-0.080	-0.033	-0.071	-0.036	-0.074	-0.038	-0.074	-0.039	-0.045	-0.036
	(0.004)	*(0.008)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.007)*	*(0.004)*	*(0.007)*	*(0.003)*	*(0.005)*
Certainty of Punishment	0.003	-0.022	0.007	-0.017	-0.001	-0.015	-0.001	-0.015	0.006	-0.019
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*
Children	-0.017	0.004	-0.015	0.008	-0.015	0.011	-0.014	0.010	-0.024	0.007
	(0.007)	*(0.008)*	*(0.008)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.008)*	*(0.008)*
Family Crime	0.004	0.150	-0.032	0.136	0.011	0.126	0.010	0.127	0.012	0.127
	(0.015)	*(0.015)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.011)*	*(0.015)*
Drug Use			0.041	0.160	-0.043	0.164	-0.044	0.164	-0.042	0.165
			(0.011)	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.009)*	*(0.011)*
Unemployment Rate	0.021	0.006	0.019	0.007	0.023	0.007	0.023	0.006	0.017	0.006
	(0.006)	*(0.006)*	*(0.006)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.006)*
Future Outlook Inventory	0.018	-0.034	0.018	-0.031	0.015	-0.029	0.015	-0.028	0.017	-0.025
	(0.011)	*(0.012)*	*(0.011)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.009)*	*(0.012)*
Years of Crime	-0.007	0.026	-0.007	0.022	-0.005	0.020	-0.005	0.020	-0.007	0.022
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.002)*	*(0.003)*
Years of Education	0.007	0.001	0.023	0.001	-0.007	0.004	-0.002	0.001		
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.005)*		
Schools per Young Person	0.331		0.183		0.324		0.325		0.127	
	(0.071)		*(0.072)*		*(0.067)*		*(0.067)*		*(0.063)*	
Lagged Enrollment	0.189		0.173		0.199		0.224			
	(0.012)		*(0.013)*		*(0.011)*		*(0.013)*			
Enrollment		0.099		0.071						

Continued on next page...

Table A.10 – continued from previous page.

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance	
	(1)		(2)		(3)		(4)		(5)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
		(0.050)		*(0.016)*						
Lagged Crime		0.156		0.147		0.144		0.159		0.141
		(0.013)		*(0.012)*		*(0.011)*		*(0.014)*		*(0.012)*
Cognitive Factor	0.043	0.061	0.038	0.019	0.008	0.029	0.005	0.030	0.006	0.035
	(0.023)	*(0.024)*	*(0.022)*	*(0.022)*	*(0.021)*	*(0.022)*	*(0.021)*	*(0.022)*	*(0.018)*	*(0.024)*
Social/Emotional Factor	0.008	-0.106	0.020	-0.071	-0.011	-0.068	-0.012	-0.067	0.008	-0.075
	(0.014)	*(0.015)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.011)*	*(0.015)*
Jail					0.100	0.091	0.375	0.008		
					(0.012)	*(0.013)*	*(0.077)*	*(0.089)*		
Enrollment (alternative)						0.056		0.048		
						(0.047)		*(0.047)*		
Years of Crime Jail							0.000	-0.001		
							(0.005)	*(0.005)*		
Years of Education Jail							-0.020	0.009		
							(0.006)	*(0.007)*		
Lagged Enrollment Jail							-0.088			
							(0.023)			
Lagged Crime Jail								-0.049		
								(0.025)		
Enrollment Jail								0.019		
								(0.026)		
Years of Education (alternative)									0.015	0.006
									(0.004)	*(0.004)*
Lagged Enrollment (alternative)									0.085	
									(0.010)	
Enrolment (alternative)										0.128
										(0.062)
Rho	-0.141		-0.083		-0.040		-0.037		-0.272	
	(0.103)		*(0.032)*		*(0.098)*		*(0.098)*		*(0.138)*	
Observations	5,232	5,232	6,236	6,236	6,236	6,236	6,236	6,236	5,139	5,139

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In column (1) we do not include drug use as an independent regressor. In column (2), enrollment is set to zero if an individual did not attend a community school. In column (3), we condition on whether the individual is interviewed in jail, and in column (4) we interact the jail dummy with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail. In column (5) enrollment is redefined as attending school for at least nine months.

Table A.11: Average Marginal Effects from Probits for Crime and Education (Violent Crime) - Robustness Checks 3 and 4

Variable	**Age-Varying Coefficients**		**Years of Crime: Quadratic**		**Years of Crime: Piecewise-linear**		**Years of Crime: Observed Experience Only**		**Contemporaneous Effect of Crime on Education – No Instrument**		**Contemporaneous Effect of Crime on Education**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.047	0.038	0.045	0.036	0.041	0.036	0.047	0.024	0.031	-0.048	0.031	0.048
	(0.021)	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.020)*	*(0.021)*	*(0.062)*	*(0.021)*	*(0.019)*
Hispanic	-0.023	-0.019	-0.023	-0.017	-0.021	-0.017	-0.022	-0.020	-0.019	-0.021	-0.019	-0.021
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*
Black	0.026	-0.031	0.027	-0.031	0.030	-0.031	0.021	-0.025	0.032	-0.029	0.032	-0.029
	(0.017)	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*	*(0.017)*	*(0.018)*
Other	0.038	-0.013	0.038	-0.014	0.040	-0.012	0.019	0.003	0.035	-0.019	0.036	-0.017
	(0.027)	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.030)*	*(0.027)*	*(0.031)*	*(0.027)*	*(0.031)*
Female	0.057	-0.077	0.056	-0.079	0.054	-0.078	0.063	-0.101	0.068	-0.075	0.067	-0.076
	(0.015)	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.014)*	*(0.016)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*
Non-intact Family	-0.049	0.012	-0.048	0.014	-0.047	0.014	-0.044	0.017	-0.048	0.009	-0.048	0.009
	(0.015)	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.014)*	*(0.016)*	*(0.014)*	*(0.016)*	*(0.014)*	*(0.016)*
Siblings	-0.002	0.004	-0.002	0.004	-0.002	0.004	-0.002	0.004	-0.003	0.004	-0.003	0.004
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*
Age	-0.087	-0.031	-0.080	-0.030	-0.081	-0.030	-0.089	-0.034	-0.066	-0.041	-0.068	-0.044
	(0.005)	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.008)*	*(0.004)*	*(0.009)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*
Certainty of Punishment	0.003	-0.019	0.003	-0.019	0.003	-0.020	0.003	-0.020	0.007	-0.019	0.007	-0.019
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Children	-0.017	0.007	-0.017	0.008	-0.018	0.008	-0.020	0.008	-0.018	0.006	-0.018	0.006
	(0.007)	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*	*(0.007)*	*(0.008)*
Family Crime	0.005	0.131	0.005	0.130	0.005	0.129	0.001	0.130	-0.029	0.130	-0.026	0.131
	(0.015)	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*
Drug Use	-0.004	0.159	-0.004	0.159	-0.005	0.159	-0.007	0.161	-0.042	0.160	-0.039	0.160
	(0.011)	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.015)*	*(0.011)*
Unemployment Rate	0.021	0.006	0.021	0.005	0.020	0.004	0.022	0.004	0.018	0.016	0.018	0.011
	(0.006)	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.006)*	*(0.005)*
Future Outlook Inventory	0.019	-0.023	0.018	-0.024	0.018	-0.024	0.022	-0.023	0.022	-0.021	0.021	-0.021
	(0.011)	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*	*(0.011)*	*(0.012)*
Years of Crime			-0.004	0.016					-0.012	0.021	-0.012	0.020
			(0.008)	*(0.008)*					*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
Years of Education			0.007	0.001	0.006	0.001	0.013	-0.003	0.006	0.003	0.006	0.003
			(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.005)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Schools per Young Person	0.327		0.332		0.336		0.280		0.287		0.295	
	(0.071)		*(0.071)*		*(0.071)*		*(0.070)*		*(0.070)*		*(0.070)*	
Lagged Enrollment			0.189		0.189		0.169		0.180		0.182	
			(0.012)		*(0.012)*		*(0.013)*		*(0.013)*		*(0.013)*	
Enrollment				0.103		0.109		0.139				
				(0.048)		*(0.048)*		*(0.054)*				
Lagged Crime				0.141		0.140		0.110		0.141		0.142
				(0.012)		*(0.012)*		*(0.014)*		*(0.012)*		*(0.012)*
Cognitive Factor	0.042	0.033	0.043	0.030	0.042	0.029	0.041	0.029	0.036	0.034	0.037	0.033
	(0.023)	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*	*(0.023)*	*(0.024)*	*(0.022)*	*(0.024)*	*(0.023)*	*(0.024)*
Social/Emotional Factor	0.006	-0.074	0.007	-0.074	0.007	-0.073	0.012	-0.073	0.024	-0.070	0.023	-0.071
	(0.014)	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.015)*	*(0.014)*	*(0.015)*
Years of Crime Age1	-0.007	0.020										
	(0.003)	*(0.004)*										
Years of Crime Age2	-0.007	0.022										
	(0.003)	*(0.003)*										

Continued on next page...

Table A.11 – continued from previous page.

Variable	**Age-Varying Coefficients**		**Years of Crime: Quadratic**		**Years of Crime: Piecewise-linear**		**Years of Crime: Observed Experience Only**		**Contemporaneous Effect of Crime on Education – No Instrument**		**Contemporaneous Effect of Crime on Education**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Education Age1	0.002 *(0.005)*	0.004 *(0.005)*										
Years of Education Age2	0.009 *(0.004)*	0.002 *(0.005)*										
Lagged Enrollment Age1	0.229 *(0.020)*											
Lagged Enrollment Age2	0.172 *(0.016)*											
Enrollment Age1		0.080 *(0.050)*										
Enrollment Age2		0.044 *(0.058)*										
Lagged Crime Age1		0.140 *(0.018)*										
Lagged Crime Age2		0.146 *(0.018)*										
Years of Crime Squared			-0.000 *(0.001)*	0.000 *(0.001)*								
Years of Crime: 0 to 4					-0.015 *(0.006)*	0.031 *(0.007)*						
Years of Crime: 5 to 9					-0.011 *(0.003)*	0.023 *(0.004)*						
Years of Crime: 10 or more					-0.005 *(0.003)*	0.025 *(0.003)*						
Years of Crime Age of Entry 14							-0.048 *(0.007)*	0.031 *(0.007)*				
Years of Crime Age of Entry 15							-0.024 *(0.006)*	0.048 *(0.006)*				
Years of Crime Age of Entry 16							-0.020 *(0.006)*	0.044 *(0.006)*				
Years of Crime Age of Entry 17							0.009 *(0.006)*	0.056 *(0.008)*				
Years of Crime Age of Entry 18							0.006 *(0.013)*	0.071 *(0.013)*				
Crime									0.202 *(0.051)*		0.188 *(0.054)*	
Lagged State Arrest Rate										-1.550 *(0.943)*		
Rho	-0.070 *(0.109)*		-0.155 *(0.104)*		-0.169 *(0.104)*		-0.237 *(0.120)*		-0.402 *(0.146)*		-0.442 *(0.141)*	
Observations	5,232	5,232	5,232	5,232	5,232	5,232	5232	5232	5232	5232	5232	5232

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In column (1), coefficients are allowed to vary by age. Age1 is a dummy for ages 14 to 19, and Age2 is a dummy for ages 20 and above. In column (2) we use a quadratic function in criminal experience. In column (3) we use a piecewise-linear function of criminal experience: 0 to 4, 5 to 9, more than 10. In column (4) we use the criminal experience observed in the sample only, interacted with age of entry dummies. In column (5) we change the direction of the contemporaneous effect; we estimate the contemporaneous effect of crime on education. In column (6) we add the lagged state arrest rate as an exclusion in the crime equation.

Table A.12: Average Marginal Effects from Probits for Crime and Education (Property Crime) - Robustness Checks 1 and 2

Variable	**Excluding Drug Use**		**Choices While in Jail (1)**		**Choices While in Jail (2)**		**Choices While in Jail (3)**		**Enrollment Based on Attendance**	
	(1)		(2)		(3)		(4)		(5)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Phoenix	0.055	0.059	0.076	0.076	0.061	0.076	0.060	0.076	0.018	0.060
	(0.021)	*(0.018)*	*(0.021)*	*(0.016)*	*(0.020)*	*(0.016)*	*(0.020)*	*(0.016)*	*(0.018)*	*(0.017)*
Hispanic	-0.023	-0.014	-0.034	-0.024	-0.005	-0.026	-0.005	-0.026	-0.021	-0.014
	(0.015)	*(0.013)*	*(0.015)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.012)*	*(0.013)*
Black	0.024	-0.010	-0.002	-0.017	0.043	-0.019	0.040	-0.019	0.002	-0.004
	(0.017)	*(0.016)*	*(0.017)*	*(0.014)*	*(0.016)*	*(0.014)*	*(0.016)*	*(0.014)*	*(0.013)*	*(0.015)*
Other	0.036	0.008	0.010	-0.010	0.034	-0.008	0.031	-0.009	-0.024	0.012
	(0.027)	*(0.025)*	*(0.027)*	*(0.024)*	*(0.026)*	*(0.024)*	*(0.026)*	*(0.024)*	*(0.022)*	*(0.025)*
Female	0.067	-0.027	0.126	-0.023	0.024	-0.012	0.025	-0.011	0.012	-0.026
	(0.014)	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.011)*	*(0.014)*
Non-intact Family	-0.048	0.007	-0.056	0.003	-0.028	0.000	-0.028	-0.001	-0.003	0.009
	(0.015)	*(0.014)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.014)*	*(0.012)*	*(0.012)*	*(0.013)*
Siblings	-0.003	0.001	-0.003	0.002	-0.002	0.002	-0.002	0.002	-0.004	0.002
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.082	-0.030	-0.072	-0.026	-0.075	-0.029	-0.075	-0.029	-0.047	-0.027
	(0.004)	*(0.006)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.003)*	*(0.004)*
Certainty of Punishment	0.004	-0.017	0.007	-0.013	-0.000	-0.013	-0.000	-0.013	0.006	-0.015
	(0.003)	*(0.002)*	*(0.003)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Children	-0.017	-0.004	-0.016	0.002	-0.016	0.003	-0.014	0.002	-0.023	-0.001
	(0.007)	*(0.007)*	*(0.008)*	*(0.006)*	*(0.007)*	*(0.006)*	*(0.007)*	*(0.006)*	*(0.008)*	*(0.007)*
Family Crime	0.000	0.112	-0.035	0.090	0.008	0.085	0.007	0.087	0.010	0.096
	(0.015)	*(0.011)*	*(0.014)*	*(0.010)*	*(0.014)*	*(0.010)*	*(0.014)*	*(0.010)*	*(0.011)*	*(0.011)*
Drug Use			0.040	0.154	-0.044	0.155	-0.045	0.154	-0.043	0.139
			(0.011)	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.009)*	*(0.010)*
Unemployment Rate	0.021	0.010	0.020	0.012	0.023	0.013	0.024	0.013	0.017	0.010
	(0.006)	*(0.005)*	*(0.006)*	*(0.004)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*
Future Outlook Inventory	0.018	-0.040	0.018	-0.037	0.014	-0.036	0.014	-0.036	0.016	-0.033
	(0.011)	*(0.010)*	*(0.011)*	*(0.009)*	*(0.010)*	*(0.009)*	*(0.010)*	*(0.009)*	*(0.009)*	*(0.010)*
Years of Crime	-0.004	0.022	-0.004	0.018	-0.002	0.018	-0.002	0.019	-0.004	0.018
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.003)*	*(0.003)*	*(0.002)*	*(0.002)*
Years of Education	0.007	0.007	0.023	0.006	-0.007	0.007	-0.002	0.007		
	(0.004)	*(0.004)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.003)*	*(0.004)*	*(0.004)*		
Schools per Young Person	0.326		0.177		0.322		0.322		0.132	
	(0.071)		*(0.072)*		*(0.067)*		*(0.067)*		*(0.064)*	
Lagged Enrollment	0.191		0.175		0.199		0.223			
	(0.012)		*(0.013)*		*(0.011)*		*(0.013)*			
Enrollment		-0.029		-0.007						

Continued on next page...

Table A.12 – continued from previous page.

Variable	Excluding Drug Use		Choices While in Jail (1)		Choices While in Jail (2)		Choices While in Jail (3)		Enrollment Based on Attendance	
	(1)		(2)		(3)		(4)		(5)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
		(0.040)		*(0.013)*						
Lagged Crime		0.163		0.142		0.140		0.156		0.145
		(0.010)		*(0.010)*		*(0.010)*		*(0.012)*		*(0.010)*
Cognitive Factor	0.040	0.049	0.035	0.007	0.006	0.012	0.003	0.013	0.004	0.014
	(0.023)	*(0.020)*	*(0.022)*	*(0.019)*	*(0.021)*	*(0.019)*	*(0.021)*	*(0.019)*	*(0.018)*	*(0.020)*
Social/Emotional Factor	0.008	-0.092	0.017	-0.065	-0.011	-0.064	-0.012	-0.064	0.009	-0.065
	(0.014)	*(0.013)*	*(0.014)*	*(0.012)*	*(0.013)*	*(0.012)*	*(0.013)*	*(0.012)*	*(0.011)*	*(0.013)*
Jail					0.100	0.045	0.388	0.061		
					(0.012)	*(0.011)*	*(0.077)*	*(0.074)*		
Enrollment (alternative)						-0.024		-0.035		
						(0.038)		*(0.038)*		
Years of Crime Jail							-0.002	-0.004		
							(0.004)	*(0.004)*		
Years of Education Jail							-0.020	0.002		
							(0.006)	*(0.006)*		
Lagged Enrollment Jail							-0.086			
							(0.023)			
Lagged Crime Jail								-0.050		
								(0.020)		
Enrollment Jail								0.019		
								(0.021)		
Years of Education (alternative)									0.015	0.005
									(0.004)	*(0.004)*
Lagged Enrollment (alternative)									0.090	
									(0.010)	
Enrolment (alternative)										-0.074
										(0.049)
Rho	0.067		0.006		0.044		0.055		0.172	
	(0.110)		*(0.036)*		*(0.107)*		*(0.107)*		*(0.147)*	
Observations	5,232	5,232	6,231	6,231	6,231	6,231	6,231	6,231	5,141	5,141

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In column (1) we do not include drug use as an independent regressor. In column (2), enrollment is set to zero if an individual did not attend a community school. In column (3), we condition on whether the individual is interviewed in jail, and in column (4) we interact the jail dummy with years of education, years of crime, and enrollment to allow the effect of previous experience and contemporaneous enrollment to vary with whether the individual is in jail. In column (5) enrollment is redefined as attending school for at least nine months.

Table A.13: Average Marginal Effects from Probits for Crime and Education (Property Crime) - Robustness Checks 3 and 4

Variable	**Age-Varying Coefficients** (1) Educ.	Crime	**Years of Crime: Quadratic** (2) Educ.	Crime	**Years of Crime: Piecewise-linear** (3) Educ.	Crime	**Years of Crime: Observed Experience Only** (4) Educ.	Crime	**Contemporaneous Effect of Crime on Education – No Instrument** (5) Educ.	Crime	**Contemporaneous Effect of Crime on Education** (6) Educ.	Crime
Phoenix	0.057	0.064	0.054	0.062	0.053	0.063	0.064	0.072	0.049	0.050	0.049	0.059
	(0.021)	*(0.017)*	*(0.021)*	*(0.017)*	*(0.021)*	*(0.017)*	*(0.021)*	*(0.018)*	*(0.022)*	*(0.054)*	*(0.022)*	*(0.016)*
Hispanic	-0.023	-0.012	-0.023	-0.013	-0.023	-0.013	-0.026	-0.020	-0.023	-0.011	-0.023	-0.011
	(0.015)	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Black	0.024	-0.005	0.025	-0.006	0.025	-0.005	0.020	-0.013	0.024	-0.006	0.024	-0.006
	(0.017)	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*	*(0.017)*	*(0.015)*
Other	0.036	0.016	0.037	0.014	0.036	0.016	0.014	0.012	0.034	0.014	0.034	0.014
	(0.027)	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*	*(0.027)*	*(0.025)*
Female	0.068	-0.022	0.067	-0.025	0.067	-0.025	0.068	-0.041	0.067	-0.027	0.067	-0.027
	(0.014)	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*	*(0.014)*
Non-intact Family	-0.049	0.007	-0.048	0.007	-0.048	0.007	-0.041	0.008	-0.049	0.008	-0.049	0.008
	(0.015)	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Siblings	-0.003	0.002	-0.003	0.002	-0.003	0.002	-0.003	0.002	-0.003	0.002	-0.003	0.002
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*
Age	-0.090	-0.026	-0.083	-0.027	-0.082	-0.027	-0.089	-0.030	-0.080	-0.024	-0.080	-0.024
	(0.005)	*(0.007)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*	*(0.007)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.003)*
Certainty of Punishment	0.003	-0.015	0.004	-0.016	0.004	-0.016	0.003	-0.015	0.004	-0.016	0.004	-0.016
	(0.003)	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Children	-0.017	0.000	-0.018	0.001	-0.017	0.001	-0.020	0.001	-0.018	0.001	-0.018	0.001
	(0.007)	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Family Crime	0.002	0.095	0.002	0.095	0.002	0.094	0.001	0.097	-0.008	0.094	-0.007	0.094
	(0.015)	*(0.011)*	*(0.015)*	*(0.011)*	*(0.015)*	*(0.011)*	*(0.014)*	*(0.011)*	*(0.017)*	*(0.011)*	*(0.017)*	*(0.011)*
Drug Use	-0.006	0.143	-0.007	0.144	-0.006	0.143	-0.005	0.145	-0.016	0.144	-0.016	0.144
	(0.011)	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.011)*	*(0.009)*	*(0.014)*	*(0.009)*	*(0.014)*	*(0.009)*
Unemployment Rate	0.022	0.011	0.021	0.010	0.021	0.010	0.022	0.010	0.021	0.009	0.021	0.009
	(0.006)	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.005)*	*(0.006)*	*(0.004)*
Future Outlook Inventory	0.019	-0.033	0.018	-0.032	0.017	-0.032	0.020	-0.030	0.019	-0.033	0.019	-0.033
	(0.011)	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*	*(0.011)*	*(0.010)*
Years of Crime			-0.011	0.025					-0.005	0.018	-0.005	0.018
			(0.006)	*(0.006)*					*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Years of Education			0.007	0.006	0.007	0.006	0.012	0.005	0.007	0.005	0.007	0.005
			(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Schools per Young Person	0.328		0.330		0.329		0.276		0.323		0.323	
	(0.072)		*(0.071)*		*(0.071)*		*(0.071)*		*(0.071)*		*(0.071)*	
Lagged Enrollment			0.190		0.190		0.175		0.191		0.191	
			(0.012)		*(0.012)*		*(0.012)*		*(0.012)*		*(0.012)*	
Enrollment				-0.022		-0.020		-0.002				
				(0.040)		*(0.040)*		*(0.044)*				
Lagged Crime				0.144		0.144		0.126		0.144		0.144
				(0.010)		*(0.010)*		*(0.011)*		*(0.010)*		*(0.010)*
Cognitive Factor	0.040	0.022	0.040	0.021	0.039	0.022	0.037	0.030	0.039	0.018	0.039	0.018
	(0.023)	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*	*(0.023)*	*(0.020)*
Social/Emotional Factor	0.005	-0.068	0.007	-0.068	0.007	-0.068	0.003	-0.076	0.012	-0.067	0.012	-0.067
	(0.014)	*(0.013)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.013)*	*(0.014)*	*(0.012)*	*(0.015)*	*(0.013)*	*(0.015)*	*(0.013)*
Years of Crime Age1	-0.004	0.024										
	(0.003)	*(0.003)*										
Years of Crime Age2	-0.003	0.014										
	(0.003)	*(0.003)*										

Continued on next page...

Table A.13 – continued from previous page.

Variable	**Age-Varying Coefficients**		**Years of Crime: Quadratic**		**Years of Crime: Piecewise-linear**		**Years of Crime: Observed Experience Only**		**Contemporaneous Effect of Crime on Education – No Instrument**		**Contemporaneous Effect of Crime on Education**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime	Educ.	Crime
Years of Education Age1	0.003 *(0.005)*	0.005 *(0.004)*										
Years of Education Age2	0.010 *(0.004)*	0.007 *(0.004)*										
Lagged Enrollment Age1	0.229 *(0.020)*											
Lagged Enrollment Age2	0.172 *(0.016)*											
Enrollment Age1		-0.035 *(0.041)*										
Enrollment Age2		-0.050 *(0.046)*										
Lagged Crime Age1		0.130 *(0.014)*										
Lagged Crime Age2		0.157 *(0.016)*										
Years of Crime Squared			0.001 *(0.001)*	-0.001 *(0.001)*								
Years of Crime: 0 to 4					-0.008 *(0.004)*	0.024 *(0.005)*						
Years of Crime: 5 to 9					-0.004 *(0.002)*	0.020 *(0.003)*						
Years of Crime: 10 or more					0.000 *(0.005)*	0.016 *(0.003)*						
Years of Crime Age of Entry 14							-0.055 *(0.009)*	0.020 *(0.007)*				
Years of Crime Age of Entry 15							-0.033 *(0.007)*	0.039 *(0.006)*				
Years of Crime Age of Entry 16							-0.022 *(0.007)*	0.036 *(0.006)*				
Years of Crime Age of Entry 17							0.008 *(0.007)*	0.040 *(0.006)*				
Years of Crime Age of Entry 18							0.021 *(0.016)*	0.050 *(0.012)*				
Crime									0.058 *(0.051)*		0.057 *(0.051)*	
Lagged State Arrest Rate										-0.137 *(0.829)*		
Rho	0.122 *(0.121)*		0.061 *(0.118)*		0.055 *(0.118)*		0.012 *(0.127)*		-0.130 *(0.121)*		-0.133 *(0.122)*	
Observations	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5,232	5232	5232	5232	5232

Notes:
1. Standard errors are reported below the point estimates in italics and in parentheses.
2. In column (1), coefficients are allowed to vary by age. Age1 is a dummy for ages 14 to 19, and Age2 is a dummy for ages 20 and above. In column (2) we use a quadratic function in criminal experience. In column (3) we use a piecewise-linear function of criminal experience: 0 to 4, 5 to 9, more than 10. In column (4) we use the criminal experience observed in the sample only, interacted with age of entry dummies. In column (5) we change the direction of the contemporaneous effect; we estimate the contemporaneous effect of crime on education. In column (6) we add the lagged state arrest rate as an exclusion in the crime equation.

Bibliography

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman. 2003. "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice." *International Economic Review* 44 (2):361–422. 2001 Lawrence R. Klein Lecture.

Cooley Fruehwirth, Jane, Salvador Navarro, and Yuya Takahashi. 2016. "Identification and Estimation of Time-varying Treatment Effects: How The Timing of Grade Retention Affects Outcomes." *Journal of Labor Economics* 34 (4). Forthcoming.

Kotlarski, Ignacy I. 1967. "On Characterizing the Gamma and Normal Distribution." *Pacific Journal of Mathematics* 20 (1):69–76.

Appendix B

Chapter 3 Appendices

Table B.1: Pathways to Desistance - Descriptive Statistics by Location

Variable	Philadelphia	Phoenix
Black*	0.698	0.093
	(0.460)	(0.291)
Hispanic*	0.145	0.581
	(0.353)	(0.494)
White*	0.133	0.287
	(0.340)	(0.453)
Age at Labour Market Entry*	18.790	18.996
	(1.229)	(1.691)
High School Degree*	0.387	0.136
	(0.488)	(0.344)
GED*	0.298	0.423
	(0.458)	(0.495)
Income Crime Monthly Rate	0.113	0.101
	(0.317)	(0.301)
Legal Employment Monthly Rate	0.261	0.401
	(0.439)	(0.490)
Monthly Non-Employment Rate	0.361	0.210
	(0.48)	(0.408)
Monthly Incarceration Rate	0.290	0.331
	(0.454)	(0.471)
Monthly Legal Earnings	1,404.8	1,440.2
	(607.9)	(490.3)
Monthly Criminal Earnings	5,240.2	3,166.5
	(5168.6)	(5629.1)
Number of Individuals	248	279

Notes:

* Indicates variables that do not vary over time. Summary statistics for these variables are calculated using only the survey at the time individuals stop attending school. For the rest of the variables, summary statistics are calculated using all individual-month observations.

1. Standard deviations are reported below the mean in parenthesis.

2. The monthly crime rate is the fraction of individual-month pairs in which an income crime is committed. The monthly legal employment rate, non-employment rate, and incarceration rate are calculated analogously.

3. Legal earnings and criminal earnings are monthly and are expressed in 2000 US dollars.

Appendix C

Chapter 4 Appendices

Table C.1: Parameter Estimates - Sensitivity Analysis - Philadelphia

	Trimming Percentages in Criminal Earnings Distribution			
Parameter	**(1) 1% Bottom 1% Top**	**(2) 1% Bottom 5% Top**	**(3) 5% Bottom 5% Top**	**(4) 5% Bottom 1% Top**
Legal Earnings Distribution: Mean	7.281	7.281	7.277	7.279
	(0.037)	(0.022)	(0.021)	(0.021)
Legal Earnings Distribution: Variance	0.240	0.240	0.241	0.241
	(0.025)	(0.014)	(0.015)	(0.015)
Criminal Earnings Distribution: Mean	8.327	8.244	8.327	8.356
	(0.183)	(0.149)	(0.168)	(0.181)
Criminal Earnings Distribution: Variance	1.401	1.237	0.989	1.258
	(0.283)	(0.149)	(0.138)	(0.194)
Jail: Release Rate	0.075	0.075	0.075	0.075
	(0.006)	(0.005)	(0.005)	(0.005)
Non-Employment: Job Arrival Rate	0.069	0.069	0.070	0.069
	(0.007)	(0.004)	(0.004)	(0.004)
Non-Employment: Crime Arrival Rate	0.037	0.037	0.037	0.038
	(0.005)	(0.004)	(0.004)	(0.004)
Non-Employment: Arrest Rate	0.034	0.034	0.034	0.034
	(0.004)	(0.003)	(0.003)	(0.003)
Legal Employment: Job Arrival Rate	0.082	0.082	0.083	0.082
	(0.015)	(0.006)	(0.006)	(0.007)
Legal Employment: Crime Arrival Rate	0.031	0.031	0.032	0.033
	(0.008)	(0.004)	(0.005)	(0.005)
Legal Employment: Job Destruction Rate	0.089	0.089	0.089	0.089
	(0.009)	(0.005)	(0.005)	(0.005)
Legal Employment: Arrest Rate	0.008	0.008	0.008	0.008
	(0.002)	(0.001)	(0.001)	(0.001)
Income Crime: Job Arrival Rate	0.027	0.025	0.026	0.027
	(0.01)	(0.005)	(0.005)	(0.006)
Income Crime: Earnings Shock Rate	0.123	0.122	0.121	0.125
	(0.021)	(0.017)	(0.017)	(0.018)
Income Crime: Destruction Rate	0.101	0.101	0.099	0.097
	(0.013)	(0.009)	(0.01)	(0.009)
Income Crime: Arrest Rate	0.095	0.096	0.095	0.095
	(0.011)	(0.007)	(0.007)	(0.007)
Employment/Crime: Job Arrival Rate	0.061	0.060	0.063	0.059
	(0.036)	(0.016)	(0.018)	(0.017)
Employment/Crime: Crime Earnings Shock Rate	0.063	0.062	0.067	0.066
	(0.031)	(0.012)	(0.013)	(0.013)
Employment/Crime: Crime Destruction Rate	0.130	0.131	0.133	0.123
	(0.038)	(0.017)	(0.018)	(0.018)
Employment/Crime: Arrest Rate	0.039	0.039	0.039	0.039
	(0.014)	(0.011)	(0.011)	(0.011)
Employment/Crime: Job Destruction Rate	0.074	0.074	0.073	0.075
	(0.027)	(0.011)	(0.01)	(0.011)
Jail: Job Offer Probability	0.086	0.086	0.084	0.085
	(0.022)	(0.01)	(0.01)	(0.01)
Jail: Crime Opportunity Probability	0.165	0.165	0.166	0.171
	(0.029)	(0.021)	(0.023)	(0.023)
Flow Utility of Jail	172.2	130.4	73.1	-54.8
	(261.7)	(100.1)	(42.2)	(-30.1)
Flow Utility of Leisure	956.1	972.2	1023.3	1095.6
	(431.8)	(284.4)	(334.5)	(289.2)
Flow Utility of Crime	674.6	751.2	577.4	702.3
	(1005.6)	(532.7)	(771.4)	(557.2)

Notes:
1. Standard errors are reported below the point estimates in parenthesis. These are computed by bootstrap with 100 replications.
2. Arrival rates are monthly.
3. In column (1), I estimate the model using a trimming percentage of 1% at the bottom and the top of the criminal earnings distribution. In column (2), I use a trimming percentage of 1% in the bottom and 5% in the top for the criminal earnings distribution. In column (3), I use a trimming percentage of 5% in the bottom and 1% in the top for the criminal earnings distribution. Finally, in column (4), I use a 5% trimming percentage in the top and the bottom of the criminal earnings distribution.
4. The flow utility of crime equals $(\alpha_c * 2/3)$, which is the non-pecuniary value of crime obtained by an individual who is only participating in income crime.

Table C.2: Parameter Estimates - Sensitivity Analysis - Phoenix

Parameter	Trimming Percentages in Criminal Earnings Distribution			
	(1) **1% Bottom** **1% Top**	**(2)** **1% Bottom** **5% Top**	**(3)** **5% Bottom** **5% Top**	**(4)** **5% Bottom** **1% Top**
Legal Earnings Distribution: Mean	7.339	7.340	7.339	7.342
	(0.024)	(0.022)	(0.022)	(0.022)
Legal Earnings Distribution: Variance	0.200	0.200	0.202	0.201
	(0.015)	(0.012)	(0.012)	(0.012)
Criminal Earnings Distribution: Mean	7.234	7.171	7.186	7.195
	(0.301)	(0.13)	(0.145)	(0.156)
Criminal Earnings Distribution: Variance	2.114	1.598	1.517	2.108
	(0.636)	(0.193)	(0.211)	(0.325)
Jail: Release Rate	0.081	0.081	0.081	0.081
	(0.008)	(0.006)	(0.006)	(0.006)
Non-Employment: Job Arrival Rate	0.102	0.102	0.102	0.102
	(0.008)	(0.006)	(0.006)	(0.006)
Non-Employment: Crime Arrival Rate	0.040	0.040	0.042	0.044
	(0.007)	(0.004)	(0.004)	(0.005)
Non-Employment: Arrest Rate	0.046	0.047	0.047	0.046
	(0.005)	(0.004)	(0.004)	(0.004)
Legal Employment: Job Arrival Rate	0.141	0.141	0.141	0.140
	(0.013)	(0.011)	(0.011)	(0.011)
Legal Employment: Crime Arrival Rate	0.031	0.031	0.032	0.031
	(0.006)	(0.004)	(0.006)	(0.005)
Legal Employment: Job Destruction Rate	0.054	0.054	0.054	0.053
	(0.004)	(0.003)	(0.003)	(0.003)
Legal Employment: Arrest Rate	0.0100	0.0100	0.0100	0.0100
	(0.002)	(0.002)	(0.002)	(0.002)
Income Crime: Job Arrival Rate	0.054	0.054	0.054	0.053
	(0.012)	(0.010)	(0.011)	(0.011)
Income Crime: Earnings Shock Rate	0.113	0.111	0.117	0.120
	(0.025)	(0.015)	(0.016)	(0.017)
Income Crime: Destruction Rate	0.132	0.136	0.128	0.123
	(0.025)	(0.012)	(0.012)	(0.012)
Income Crime: Arrest Rate	0.109	0.109	0.110	0.109
	(0.014)	(0.008)	(0.008)	(0.008)
Employment/Crime: Job Arrival Rate	0.047	0.047	0.045	0.043
	(0.015)	(0.013)	(0.013)	(0.012)
Employment/Crime: Crime Earnings Shock Rate	0.102	0.103	0.102	0.099
	(0.019)	(0.021)	(0.019)	(0.020)
Employment/Crime: Crime Destruction Rate	0.172	0.173	0.169	0.168
	(0.034)	(0.022)	(0.023)	(0.024)
Employment/Crime: Arrest Rate	0.018	0.018	0.017	0.017
	(0.006)	(0.005)	(0.005)	(0.005)
Employment/Crime: Job Destruction Rate	0.071	0.071	0.073	0.075
	(0.01)	(0.011)	(0.01)	(0.011)
Jail: Job Offer Probability	0.234	0.236	0.236	0.237
	(0.036)	(0.026)	(0.027)	(0.027)
Jail: Crime Opportunity Probability	0.208	0.206	0.218	0.225
	(0.048)	(0.026)	(0.03)	(0.031)
Flow Utility of Jail	373.6	553.7	613.8	206.4
	(317.4)	(425.3)	(354.3)	(113.4)
Flow Utility of Leisure	1433.8	1316.3	1262.9	1537.6
	(326)	(385.1)	(412.8)	(405.9)
Flow Utility of Crime	1539.3	1355.2	1194.4	1675.2
	(581.1)	(961.1)	(1595.8)	(1329.1)

Notes:
1. Standard errors are reported below the point estimates in parenthesis. These are computed by bootstrap with 100 replications.
2. Arrival rates are monthly.
3. In column (1), I estimate the model using a trimming percentage of 1% at the bottom and the top of the criminal earnings distribution. In column (2), I use a trimming percentage of 1% in the bottom and 5% in the top for the criminal earnings distribution. In column (3), I use a trimming percentage of 5% in the bottom and 1% in the top for the criminal earnings distribution. Finally, in column (4), I use a 5% trimming percentage in the top and the bottom of the criminal earnings distribution.
4. The flow utility of crime equals $(\alpha_c * 2/3)$, which is the non-pecuniary value of crime obtained by an individual who is only participating in income crime.

Curriculum Vitae

Name:	Maria Antonella Mancino
Post-Secondary Education and Degrees:	The University of Western Ontario London, Ontario, Canada 2012 - 2018 (Expected), Ph.D. in Economics
	Universidad del CEMA Buenos Aires, Argentina 2009-2010, M.A. in Economics
	Universidad Nacional de Cuyo Mendoza, Argentina 2004-2008, B.A. in Economics
Honours and Awards:	Western Graduate Research Scholarship The University of Western Ontario 2012 - 2017
	Sir Arthur Currie Memorial Scholarship The University of Western Ontario 2013
	Graduate Teaching Assistant of the Year The University of Western Ontario 2014
	Summer Paper Prize The University of Western Ontario 2015
Related Work Experience:	Teaching Assistant The University of Western Ontario 2012 - 2017

Research Assistant
The University of Western Ontario
2013 - 2018

Publications ”Separating State Dependence, Experience, and Heterogeneity in a Model of Youth Crime and Education”, with Salvador Navarro and David Rivers, Economics of Education Review 54: 274-285, 2016.