3-3-2020 1:00 PM

# Essays on Criminal Behaviour, Human Capital Formation, and Mental Health

Diego F. Salazar, *The University of Western Ontario*

Supervisor: Lochner, Lance J., *The University of Western Ontario*
Joint Supervisor: Navarro, Salvador, *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 contribute to the study of some of the negative consequences of incarceration and their relation with the life-cycle choices of juvenile offenders.

Chapter 2 studies the causal relationship between incarceration and mental health problems. In this chapter, I use different matching estimators to identify the causal effects of incarceration on several dimensions of mental health using data from a survey of juvenile offenders, the Pathways to Desistance (PTD) survey. My findings show that being incarcerated for the first time, between 17 and 18 years old, increases depression by at least 0.18 standard deviations and hostility by at least 0.17 standard deviations. I also find evidence of heterogeneous effects for depression, hostility, and an overall measure of mental health, with bigger impacts for people with a low incarceration probability and blacks. I calculate the relation between the deterioration of mental health and changes in criminal activity during the year following the treatment period. My results show that depression is associated with an increase of at least 1.62 percentage points in crime, while hostility is associated with an increase of at least 1.19 percentage points.

Chapter 3 documents empirical facts about the dynamic interplay between choices and personal capabilities. I estimate the relationship between mental health and self-control, criminal engagement, and human capital accumulation, for the period that spans the transition from adolescence into adulthood. I take advantage of the PTD survey and control for time-invariant heterogeneity and an extensive set of covariates that have been absent from the majority of previous studies about the continuity in criminal behaviour. My results suggest that self-control is an important predictor for high school graduation and that mental health problems have a negative association with high school graduation and a positive one with criminal participation. I find strong evidence that personal capabilities are malleable. I provide evidence that the experience of incarceration and the accumulation of criminal experience are associated with higher levels of depression, while the accumulation of experience in the legal sector is associated with improvements over different dimensions of mental health. I find evidence that high school graduation and employment influence the evolution of self-control. My results also provide evidence about the returns to criminal experience and incarceration records in the illegal sector.

Motivated by the findings from the previous chapters, Chapter 4 analyses the short- and long-run choices of juvenile offenders. For this purpose, I develop and estimate a dynamic model of employment, schooling, and crime with endogenous human and criminal capital that incorporates the evolution of mental health and self-control. I use this model to gain insight into the ways in which the dynamic interaction between human and criminal capital influences the life-cycle choices of juvenile offenders. My results indicate that criminal capital accumulates at a faster rate than human capital, which reduces the deterrent effects of detention, education, and employment. Further, years of schooling make the most significant contribution to the accumulation of human capital, and lower self-control is associated with incarceration, criminal engagement and unemployment. I also discuss the effects of school and wage subsidies on employment and criminal engagement and find that, depending on the age at the intervention, both policies can generate long-run crime reduction.



**Keywords:** Crime, education, employment, health, human capital, mental health.

# Summary for Lay Audience

My thesis consists of three chapters that contribute to the study of some of the negative consequences of incarceration and their relation with the life-cycle choices of juvenile offenders.

Chapter 2 studies the causal relationship between incarceration and mental health problems. My findings show that being incarcerated for the first time, between 17 and 18 years old, increases depression by at least 0.18 standard deviations and hostility by at least 0.17 standard deviations. I also find evidence of heterogeneous effects for depression, hostility, and the overall measure of mental health, with bigger impacts on people with low incarceration probability and blacks. I calculate the relation between the deterioration of mental health and changes in criminal activity during the year that followed the treatment period. My results show that depression is associated with an increase of at least 1.62 percentage points in crime while hostility is associated with an increase of at least 1.19 percentage points.

Chapter 3 documents empirical facts about the dynamic interplay between choices and personal capabilities. I estimate the relationship between mental health and self-control, criminal engagement, and human capital accumulation, for the period that spans the transition from adolescence into adulthood. I take advantage of the PTD survey and control for fixed unobserved heterogeneity and an extensive set of covariates that have been absent in the majority of previous studies about the continuity in criminal behaviour. My results suggest that self-control is an important predictor for high school graduation and that mental health problems have a negative association with high school graduation and a positive one with criminal participation. I find strong evidence that personal capabilities are malleable. I provide evidence that the experience of incarceration and the accumulation of criminal experience are associated with higher levels of depression, while the accumulation of experience in the legal sector is associated with improvements over different dimensions of mental health. I find evidence that high school graduation and employment influence the evolution of self-control. My results provide evidence about the returns to criminal experience and incarceration records in the illegal sector.

Motivated by the findings from the previous chapters, Chapter 4 analyses the short- and long-run choices of juvenile offenders. For this purpose, I develop and estimate a dynamic model of employment, schooling, and crime with endogenous human and criminal capital that incorporates the evolution of mental health and self-control. I use this model to gain insight into the ways in which the dynamic interaction between human and criminal capital influences the life-cycle choices of juvenile offenders. My results indicate that criminal capital accumulates at a faster rate than human capital, which reduces the deterrent effects of detention, education, and employment. Further, years of schooling make the most significant contribution to the accumulation of human capital, and lower self-control is associated with incarceration, criminal engagement and unemployment. I also discuss the effects of school and wage subsidies on employment and criminal engagement and find that, depending on the age of the intervention, both policies can generate long-run crime reduction.

**To Lucia, Tomas, and Matias**

# Acknowlegements

I am greatly indebted to my supervisors Salvador Navarro and Lance Lochner for their guidance in developing this project. I am also grateful to David Rivers and Nirav Mehta for taking the time to talk and help with my research.

I would also like to extend my gratitude to my fellow graduate students: Antonella Mancino, Miguel Cardoso, Qian Liu, Tom Handler, Galyna Grynkiv, Sergii Pypko, Martin Luccioni, Cecilia Diaz, Francisco Adame, Aldo Sandoval, and Ali Kamranzadeh for being a permanent source of feedback and for the memories we built together.

I am grateful to my wife Lucia. She, with absolute devotion, was by my side during these years in the program sharing all the highs and lows. Most importantly, she always makes sure I am ready for the next round no matter how hard was the previous one. Finally, I would like to thank my father, mother, and brother for supporting all my aspirations and encouraging me to follow my passion.

# Contents

# List of Figures

# List of Tables

# List of Appendices

# Chapter 1

# Introduction

This thesis explores some of the negative consequences of incarceration and their relation with the life-cycle choices of juvenile offenders. I also evaluate if policies designed to affect the opportunity cost of crime, like wage and school subsidies, have the potential to reduce crime and increase employment over a long period. Both purposes are motivated by the need to find "alternative" crime-fighting strategies with the potential of reducing crime in the short- and long-run while reducing the number of people in detention.

Several reasons justify this need. First, there is ample evidence documenting the positive effect of incarceration on reoffending. Second, incarceration is an expensive policy intervention. Third, there is documented evidence of incarceration having negative consequences on the re-socialization process of former inmates.

With these reasons in mind, Chapter 2 focuses on analyzing one of the side effects of incarceration: mental health problems. This chapter provides evidence of the causal link between mental health and incarceration. It also illustrates how these changes in mental health affect the probability of future criminal engagement.

Chapter 3 provides stylized facts about the interplay between personal capabilities, defined as mental health, cognitive and noncognitive skills, and choices. It documents the association between noncognitive skills and mental health, and decisions like high school graduation, employment and crime. It also evaluates if previous choices, not only incarceration, have the potential of shaping personal capabilities. Finally, it presents evidence on how previous choices affect the continuation of criminal engagement through changes in illegal earnings and wages. The analysis in this chapter is purely descriptive and provides evidence that supports the modelling choices made for the dynamic model developed and estimated in Chapter 4.

Based on the findings from the previous chapters, Chapter 4 studies the influence of human and criminal capital on the decisions of juvenile offenders. For this purpose, I develop and estimate a life-cycle model of employment, schooling and crime from ages 13 to 32. This model allows for the endogenous evolution of both types of capital based on choices made by individuals. I incorporate multiple personal capabilities (mental health, noncognitive skills and cognitive skills) that influence the evolution of human and criminal capital, and affect the rewards available from schooling and unemployment. Furthermore, I allow for previous choices to affect the evolution of mental health and noncognitive skills. In this way, I can explore the quantitative implications of personal capabilities and criminal capital changing in response to decisions and the experience of incarceration, which has not been done in previous studies. With my estimates in hand, I analyze the short- and long-run deterrent effects of wage and school subsidies.

For the analysis in all chapters, I use a rich panel dataset on serious juvenile offenders, the Pathways to Desistance (PTD) survey. PTD is a multi-site, longitudinal dataset of serious adolescent offenders as they transition from adolescence into early adulthood. The study follows 1,354 adolescents who were found guilty of a serious criminal offence, predominantly felonies.[1] The enrolment took placed between November 2000 and January 2003 in the juvenile and adult court systems of Maricopa County (Phoenix) and Philadelphia County (Philadelphia).[2] All participants were at least 14 years old, and under 18 years old at the time they committed the offence for which they were called in to participate.

The initial survey, or baseline interview, occurred when respondents first entered the sample, which happened within 75 days of adjudication for youths in the juvenile system and, for those in the adult system, within 90 days of either a decertification hearing in Pennsylvania or an adult arraignment hearing in Arizona. After that, there were six semi-annual follow-up interviews, followed by four annual interviews.

Interviews were done at the participants' home, institutional placement, or in a public place such as a library, and took around two hours to complete. Participants were paid on a graduated payment scale designed to encourage continued participation. Payment began at $50 per interview and was capped at $150. In total, the survey followed respondents for 84 months, and the retention rate was 84% of the sample by the last follow-up interview (Mulvey et al., 2014).

PTD was designed specifically to study the evolution of criminal behaviour during the period that covers the transition from adolescence into adulthood. It has comprehensive and repeated measures of health, mental health, personality traits, and noncognitive skills. It has data about family background and the criminal history of the respondents. In addition, it contains a rich panel of information about decisions to participate in crime (and illegal earnings), have legal employment (and wages), and enroll in school (and academic achievements). These features make the PTD data well-suited for understanding the dynamics of crime, employment, and human capital formation.

There are four fundamental findings from my analysis. First, incarceration has negative consequences on the mental health of inmates. Such effects manifest mainly through increases in depression and hostility. Also, I find evidence of heterogeneous effects on depression, hostility, and the overall measure of mental health, with bigger impacts on people with low incarceration probability and blacks.

Second, there is evidence of the dynamic interplay between personal capabilities and choices. My results show this interplay is particularly strong for crime and schooling choices. In particular, high school graduation both influences and is influenced by self-control. In this case, having more noncognitive skills contributes to the formation of human capital, but also higher levels of human capital help to develop noncognitive skills. The existence of mental health problems increases the likelihood of participating in crime, but also criminal engagement (and incarceration) has a negative influence over mental health.

[1] The study capped the proportion of male juveniles with drug offences to 15% of the sample at each site to guarantee some heterogeneity.

[2] The sample represents approximately one in three adolescents adjudicated on charges in each locale during the recruitment period (Loughran et al., 2013).

Third, the dynamic interaction between criminal and human capital influences the life-cycle choices of juvenile offenders since: (i) during the first periods of sector-specific experience, criminal capital accumulates at a faster rate than human capital, (ii) during periods of incarceration, when the evolution of human capital is restricted, criminal capital continues to grow either through the accumulation of criminal experience or through the criminogenic effect of detention, (iii) higher levels of self-control are associated with lower illegal earnings which in turn is associated with less criminal engagement and fewer people in detention in the long-run, and (iv) schooling makes the most significant contribution to human capital accumulation and prevents the accumulation of criminal capital.

Fourth, wage and school subsidies have the potential to reduce crime and increase employment during the policy period (which I refer to as the short-run). However, the timing of the policies are critical to determine if they have the potential to cause the same effects after the policy period (which I refer to as the long-run). Using my estimates from the dynamic model, I show two things. The first one is that school subsidies given to adolescents generate investments in education that increase human capital and prevent the accumulation of criminal capital. Since incarceration makes a positive contribution to criminal capital, the deterrent effect of school attendance is also amplified by preventing people from being incarcerated. Over the long-run, this policy increases employment and decreases criminal engagement.

Second, wage subsidies, targeted at young adults, increase the number of economically self-sufficient people and reduce criminal engagement over the long-run. Interestingly, this policy also increases investments in education, even before the start of the subsidy period. Since expected wages are higher due to the subsidy, forward-looking agents have an incentive to remain in school and reduce their criminal engagement to increase their probability of employment during the subsidy period.

# Chapter 2

# Heterogeneous Effects of Incarceration on Mental Health

## 2.1 Introduction

The United States has the highest incarceration rate among OECD countries. What makes the United States different from other countries is the punitiveness of its criminal justice policies: the ratio of those incarcerated to those convicted is 70% higher in the United States than the next highest country. This high rate of incarceration also extends to juveniles. In 2010, the stock of juvenile detainees was over 70,000, a rate of 2.3 per 1,000 aged 10-19 (Aizer and Doyle, 2015). Including those under correctional supervision, the United States has a juvenile corrections rate that is five times higher than the next highest country (Hazel, 2008). Given the large population of inmates and ex-convicts, and the amount of public expending devoted to expanding and improving the penitentiary system, the effects of incarceration have been a hot topic of debate.

Research on the collateral consequences of incarceration has grown considerably in recent years. Usually, economists have focused on the direct consequences over educational attainment (Aizer and Doyle, 2015), employment and earnings (Western et al., 2001; Harding et al., 2018; Western, 2002), and reoffending (Mueller-Smith, 2015). However, there is also abundant evidence about the side-effects of incarceration on things like the racial gap in AIDS (Johnson and Raphael, 2009), divorce (Western, 2006; Massoglia et al., 2011), child well-being (Wakefield and Wildeman, 2011), health (Wildeman, 2012; Massoglia, 2008; Massoglia and Pridemore, 2015; Porter, 2014; Barnert et al., 2016), and mortality (Binswanger et al., 2007; Farrell and Marsden, 2008; Pridemore, 2014).

The documented side-effects of incarceration also include negative consequences over mental health. Incarceration has been linked to a range of psychological problems like anxiety and mood disorders (Fazel and Danesh, 2002; Haney, 2017; Schnittker and John, 2007; Schnittker et al., 2012; Fagan and Kupchik, 2011). Some studies even suggest that these problems are long-lasting and develop over the life course of former inmates (Baćak et al., 2019).

A different stream of literature has provided abundant evidence about the relationship between mental health problems, especially depression, and bad economic outcomes. Studies have suggested that mental health problems are especially detrimental for people who are in the early stages of development (i.e. children and adolescents). First, depression during adolescence is correlated with lower academic achievement and noncognitive development (Cook et al., 2009). Second, studies estimate that half of the adults who suffer from mental health issues had symptoms that began in adolescence (Dick and Ferguson, 2015). Third, mental health problems during adolescence impact labour market outcomes during adulthood (Fletcher, 2012). Fourth, the economic costs associated with mental health problems are substantial. Between 1996 and 2006, mental health expenditure rose rapidly from $35.2 to $57.5 billion and from the fifth to the third most costly medical condition in the United States (Agency for Healthcare Research and Quality, 2014).

Some of these problems could be even more problematic among disadvantaged young individuals with a high propensity for criminal behaviour. The observed correlation between mental health problems, crime, and lower human capital accumulation suggests that mental health impacts criminal behaviour in at least two ways. First, mental health problems are associated with higher criminal propensity (Frank and McGuire, 2010). Second, the lower human capital associated with mental health problems reduces the opportunity cost of crime.

In this way, incarceration is determined to create feedback over reoffending through mental health. Given the size of the penitentiary system under consideration, such feedback is expected to have sizable consequences over aggregate crime rates. However, there is not much evidence measuring the causal effect of incarceration on mental health, how this effect changes across different groups of the population, and its impact on reoffending.

This chapter presents two main contributions. First, it provides credible evidence for incarceration causing mental health problems. The key challenge with establishing a causal effect of incarceration is the issue of endogeneity coming from unobservable characteristics governing selection into incarceration and the evolution of mental health. In my context, it could be that unobservable characteristics from people with mental health problems are correlated with criminal behaviour and incarceration. Further, it could be that people choose crime and incarceration as a result of negative shocks to mental health. Many of the studies previously mentioned (Fagan and Kupchik, 2011; Schnittker and John, 2007; Schnittker et al., 2012) derived their conclusions based on regression analysis that identified the effects of incarceration based on the "selection on observables" assumption (Heckman and Robb, 1985). To the extent that controls used are insufficient to account for the selection into incarceration, the results from those papers would still suffer from selection bias.

To address this problem, I use a rich data set, PTD survey, to estimate the causal effect using different nonparametric matching algorithms. With the PTD in hand, I can address the endogeneity problem in two different ways. First, PTD provides information about the critical variables that simultaneously influence incarceration and mental health. In particular, there is data on key elements considered by judges to determine whether an offender should be placed in a residential facility such as (i) criminal history, (ii) the offence level of the crime, and (iii) personal characteristics used to determine the offender's threat to society and reoffending probability. For each participant, there is information on contacts with the judicial system before the start of the survey and age for the criminal onset, which are critical for establishing the person's criminal history. Also, there is data on the severity of the crime that brought the respondent into the survey which provides information on the offence level. Finally, there is information on the personal characteristics used by judges to evaluate the risk of reoffending including race, age, parents' criminal history, education level, employment status, mental health condition,[1] and drug and alcohol abuse history.[2] This information has been absent in the studies previously mentioned. Second, given the longitudinal nature of the survey, I can also control for unobserved fixed characteristics.

The second contribution is to present evidence about the existence of heterogeneous effects of incarceration on mental health. This finding has important policy implications. To design cost-effective policies, it is crucial to have information about who is suffering the worst consequences of incarceration.

To illustrate the changes in mental health caused by incarceration, I calculate the effect of having mental health problems on the probability of criminal engagement. These findings open the door to propose new crime-fighting strategies to complement law enforcement policies. In particular, public health policies designed to improve the mental health of inmates and the population with the highest risk of criminal engagement have the potential to reduce reoffending.

For the analysis, I work with the Global Severity Index (GSI) of mental health problems as measured by the Brief Symptom Inventory (BSI) test. In addition to the aggregate index, I analyze the effects of incarceration over four particular dimensions of mental health: depression, anxiety, somatization, and hostility. These are the four factors that repeatedly emerge across most factor analytic studies regarding mental health (Skeem et al., 2006).

To estimate the effects of incarceration, I first estimate the probability that an adolescent from my sample has his/her first experience with incarceration between 17 and 18 years old.[3] I use only the first experience to isolate the fact that impacts could change with the number of times the individual has been incarcerated. Given that I do not have information about the complete incarceration history before the start of the survey and many of the participants have a record of contacts with the penitentiary system, it is not possible to disentangle how the previous experiences affect the current mental health status.

I use the predicted probabilities to produce matching estimates using different matching algorithms. After that, the existence of heterogeneous treatment effects is analyzed. First, I show how effects vary with the incarceration probability. Second, I discuss how the Average Treatment Effects (ATE) vary across racial groups and gender. With these estimates in hand, I calculate the potential impact of mental health on criminal engagement during the year that follows the treatment period.

Based on the analysis, I reach three important conclusions. First, incarceration deteriorates two dimensions of mental health: depression and hostility. According to my estimates, the ATE of incarceration is an increase of at least 0.18 standard deviations for depression and 0.17 standard deviations for hostility. There is no evidence of significant effects on anxiety and somatization.

[1]There is data on the respondent's mental health condition since the beginning of the survey. Jointly with information on drug and alcohol abuse, they are fundamental predictors for the evolution of mental health.

[2]In fact, these characteristics are included in most of the algorithms currently used to predict future criminal behaviour. Information coming from those algorithms is used by judges for pretrial risk assessments (Kehl and Kessler, 2017). For example, they are mentioned in a document to inform members of the Maricopa Superior Court, Arizona (one of the locations in the PTD survey) about the benefits of using such predictive algorithms (Knox, 2017).

[3]Therefore, my subsample for estimation is composed of all of those with no incarceration experience before age 17.

Second, when I use only the “thick” support region to identify the effects of incarceration, my estimates are larger for depression and sometimes for hostility. Two things are consistent with such findings: (i) the existence of heterogeneous treatment effects with a bigger impact on the middle values of the propensity score and (ii) lingering selection on unobservables has a smaller effect for values outside the “thick” support. For depression, my results show that effects in the “thick” support are bigger than in the full support even after controlling for fixed unobservable characteristics. This is compatible with the existence of heterogeneous effects. The existence of heterogeneous effects is then confirmed when I analyze how effects change with incarceration probability, race and gender. I find evidence that for overall mental health (i.e. GSI) and especially depression, people with the lowest probability of incarceration suffer the worst consequences. Also, people with the highest probability of going into detention experience an improvement in GSI and depression.[4] Regarding hostility, I find evidence that effects in the middle values of the incarceration probability distribution are higher than in the full support. Finally, my estimates show that blacks are suffering the worst consequences of incarceration.

Third, the deterioration of mental health is related to increases in criminal participation between ages 18 and 19. According to my estimates, an increase in depression and hostility, equivalent to the one caused by incarceration, is associated with an increase of at least 2.81 percentage points in the probability of committing crime.

The remainder of this chapter proceeds as follows. Section 2.2 describes the PTD data I use in this chapter. Section 2.3 defines my parameters of interest, the identification strategy and the empirical approach I use to estimate the impacts of incarceration on mental health. Section 2.4 presents my estimates. Section 2.5 discusses how the estimates change in the “thick” support region of the propensity scores. Section 2.6 shows the incarceration effects for blacks and males. Section 2.7 presents evidence on how the effects of incarceration vary with the propensity score. Section 2.8 discusses the implications of changes in mental health on the propensity of subsequent (i.e. after the treatment) criminal participation. Finally, Section 2.9 concludes.

## 2.2 Data

### 2.2.1 PTD Survey

For the analysis, I use data from the PTD survey. The baseline and the follow-up interviews collected comprehensive measures of mental health. For this study, I focus on a subset of those measures given by the BSI. The BSI is a 53-item self-report inventory in which participants rate the extent to which they have been bothered in the past week by various symptoms.[5] The BSI has different subscales designed to assess individual symptom groups. According to the BSI scale system, bigger values are associated with the existence of mental health problems. In other words, lower values mean better mental health.[6] Four factors repeatedly emerge across

[4] By improvement I mean smaller depression and GSI.

[5] The values for the original scale are 0 =”not at all” to 4=”extremely”.

[6] The subscales are: (i). Somatization (distress arising from perceptions of bodily dysfunction), (ii) Obsession-Compulsion (thoughts and impulses that are experienced as unremitting and irresistible but are of an unwanted

most factor analytic studies based on the BSI: depression, anxiety, somatization, and hostility (Skeem et al., 2006). These are the subscales I use in the present study. I also use the global severity index (GSI) which is calculated by averaging all answers in the BSI.

The BSI measures have been used in previous studies analyzing the link between mental health and crime (Frank and McGuire, 2010), incarceration (Fagan and Kupchik, 2011), low educational attainment (Reynolds et al., 2007), and unemployment (Ettner et al., 1997). Throughout this dissertation, I work with a standardized version of these measures.[7]

Respondents also provided information about their family background and composition, criminal history and records,[8] employment, academic achievement, enrolment in school, detention status (freedom vs. detention), and the facility type where they served time.[9] During the follow-up interviews, each participant completed different monthly calendars covering the period between the current and the last interview, i.e. recall period. These calendars have monthly information about money earned from illegal activities,[10] criminal engagement, wages, and school enrolment. During the follow-up interviews, information regarding drug and alcohol abuse was also collected.

The information contained in the calendars was self-reported. Except for detention status, it was not subject to external validation.[11] To encourage accurate self-reporting, responses are kept confidential, and participants received a certificate of confidentiality from the U.S. Department of Justice. This guarantees that offences declared in the calendars would not be prosecuted based on their responses.

### 2.2.2 Definition of the Treatment and Subsample for Estimation

I compare four different dimensions of mental health (i.e. depression, anxiety, somatization, and hostility) and the GSI for individuals who have experienced incarceration for the first time

---

nature), (iii) Interpersonal Sensitivity (feelings of personal inadequacy and inferiority in comparison with others), (iv) Depression (symptoms of dysphoric mood and affect as well as lack of motivation and loss of interest in life), (v) Anxiety (nervousness and tension as well as panic attacks and feelings of terror), (vi) Hostility (thoughts, feelings or actions that are characteristic of anger), (vii) Phobic anxiety (persistent fear response to a specific place, object or situation that is irrational), (viii) Paranoid ideation (disordered thinking characteristic of projective thoughts, hostility, suspiciousness, grandiosity, fear of loss of autonomy, and delusions) and (ix) Psychoticism (withdrawn, isolated, schizoid lifestyle as well as first-rank symptoms of schizophrenia such as thought control).

[7] A standardized variable (sometimes called a z-score or a standard score) is a variable that has been rescaled to have a mean of zero and a standard deviation of one. For a standardized dimension of mental health, each case's value on the standardized variable indicates difference from the mean of the original variable in the number of standard deviations (of the original variable). The mean and standard deviation is calculated from all the valid responses in PTD.

[8] For this study, I use information on the age of the first offence.

[9] The categories for the facility type are (i) Adult facility, (ii) Juvenile facility, and (iii) Specialized facility. For this study, I focus on the first two categories.

[10] The exact wording of the question is: "You mentioned that you had made money during the past N months from ways besides working. Did you make any money during this month from activities that are illegal?". In case the answer was affirmative, additional questions were presented like the months when the antisocial activities took place, the type of crimes committed, the number of weeks worked, and the weekly money earned.

[11] Knowledge of a residential stay was obtained from multiple sources: (a) information obtained in the course of trying to locate a subject for a time-point interview, (b) information provided during a time-point interview, as reported on a revised version of the Child and Adolescent Services Assessment, and/or (c) a check, done regularly, of the Department of Corrections websites and parole hearing schedules and calls to local jails.

in their lives when they were between 17 and 18 years old with those who have not. The control group is defined as all the participants in the PTD survey who did not experience incarceration up to age 18. The treatment group is defined by those who experienced incarceration for the first time when they were 17 to 18 years old. I adopted this definition for two reasons. First, the sample size for this age group is the biggest. Second, by comparing people of the same age, I adjust for the documented correlation between age and mental health (Layard, 2017; Frank and McGuire, 2010). I define incarceration as spending time in a residential facility for more than 8 days during the recall period. For this chapter, incarceration and detention are interchangeable terms.

Table A.1 contains descriptive statistics for the full sample and the sample I used for estimation in terms of the variables used in the propensity score (See Section 2.3.2).[12]

For the full sample: (i) 49% of participants did not have, before the baseline interview, incarceration records and (ii) even though all subjects were found guilty and convicted of a serious offence, only 52% were sanctioned with detention. The subsample for estimation has 303 people with 161 in the control group and 142 in the treatment group. From the subsample, 12% served time in a facility that belongs to the adult-system and 31% in a juvenile facility.[13]

## 2.3 Empirical Strategy

### 2.3.1 Identification

Let $Y_1$ be the outcome in the "treated" state and $Y_0$ be the outcome in the "untreated" state. In this chapter, one group received the treatment (i.e. incarceration for the first time between ages 17 and 18), and one group was not in a residential placement up to age 18. Therefore, $Y_1$ corresponds to the potential outcome associated with incarceration and $Y_0$ corresponds to the potential outcome associated with freedom. These are called potential outcomes because only one of $(Y_1, Y_0)$ can be observed for each person. Let $D = 1$ indicate that a person was incarcerated when he/she was 17-18 years old (i.e. received the treatment) and $D = 0$ indicates that a person did not receive the treatment. Finally, let $X$ be a vector of observed characteristics affecting both the probability of being incarcerated and the outcomes of interest.

The parameter of interest is the ATE. In terms of the previous notation, the parameter is:

$$\Delta^{ATE} = \mathbb{E}(Y_1 - Y_0). \tag{2.1}$$

[12]The first two columns present the mean and standard deviation for the full sample. Columns three and four present the same descriptive statistics for the subsample for estimation. Finally, the last column presents the p-value associated with the $t$ tests on the equality of means between the two groups.

[13]The demographic characteristics in the full sample and the subsample for estimation are similar. In both cases, the average age of the participants at the baseline interview was around 16 years old, the average age of the first (self-reported) offence was a little bit over 10 years old, and the average age of the first official petition in the criminal justice system was close to 15 years old. In both samples, there is an over-representation of minorities (over 75%), and both locations are evenly represented. Regarding the sanction associated with the petition that brought participants into the survey, in the full sample, the majority (51%) was placed in a residential facility and 42% went on probation. For the subsample, the majority went on probation (60%) and 35% went into incarceration. In section 2.3.2, I provide additional details about the subsample for estimation.

There are important differences between the two groups in terms of the proportion of whites, hispanics and males, family background, cognitive ability, and the proxy for previous criminal experience.

In this chapter, I use different matching algorithms to estimate the parameter of interest. These include nearest neighbour matching, Inverse Probability Weighting (IPW), Inverse Probability Weighting with Regression Adjustment (IPWRA), and Matching combined with Difference-In-Differences (MDID).

To identify $\Delta^{ATE}$ matching techniques assume that the set of observables, $X$, contains all the information about the potential outcomes in the absence of the treatment, $Y_0$ and $Y_1$, that was used to decide the treatment, $D$. In this way, conditioning on an available set of covariates removes all systematic differences in outcomes in the "untreated" state between people who were sent to incarceration and people who were not. The "selection on observables" assumption that justifies matching is normally referred to as the *Conditional Independence Assumption* (CIA) and it is stated as follows:

$$(Y_0, Y_1 \perp D) \mid X. \tag{CIA-ATE}$$

This assumption states that, conditional on $X$, there are no unobservable elements that affect the outcomes and participation in the treatment at the same time. In this way, cross-sectional matching accounts only for selection on observables. Moreover, matching methods rely on the *common support* assumption, which can be expressed as

$$0 < Pr(D = 1 \mid X) < 1. \tag{2.2}$$

The *common support* condition states that, for each $X$ satisfying *CIA*, there must be some individuals who do not get treated. In the context of this paper, this means that, for each $X$ for which somebody had experienced incarceration, there must be someone who was not placed in a residential facility.

A serious limitation to the implementation of matching is the dimensionality of the space of the matching variables, $X$. With high-dimensional $X$, the number of different vector values becomes large, and many of the treated persons will have no counterpart in the control group. In the same way, many of the controls will have no counterpart in the treatment group.

The most common solution to this problem is to reduce the dimension of $X$ by matching on the probability of treatment, $P(X) = Pr(D = 1 \mid X)$, which is usually referred as "propensity score matching" (Rosenbaum and Rubin, 1983). The intuition behind propensity score matching is that subgroups, with values of $X$ that imply the same probability of treatment, can be matched. This is because they will appear in the treatment and (matched) control groups in the same proportion. As a consequence, matching balances the distribution of all relevant pre-treatment characteristics, $X$, in the treatment and comparison group. In this way, it achieves independence between the potential outcomes and the assignment into treatment. To estimate ATE, the counterfactuals $E(Y_0 \mid D = 1)$ and $E(Y_1 \mid D = 0)$ can be approximated using the control group and the treatment group respectively if

$$(Y_0, Y_1 \perp D) \mid P(X). \tag{CIA-ATE'}$$

To deal with potential selection on time-invariant unobservables (Heckman et al., 1997), I exploit the longitudinal nature of the data and use MDID. For MDID, I also assume that, conditional on observables $X$, the evolution of the unobservable part of the outcome is independent of the treatment status. Define $u_{t_1}$ to be the unobserved part of the outcome after the treatment and $u_{t_0}$ to be the unobserved part before the treatment. Assuming additive separability between the observed and unobserved components, the *CIA* in the context of MDID can be stated as follows:

$$(u_{t_1} - u_{t_0} \perp D) \mid P(X). \tag{CIA-MDID}$$

### 2.3.2 The Propensity Score

The estimation is done in two steps. First, I estimate the propensity score using a binary logistic model with incarceration as the dependent variable. Second, I calculate the weighted average of the difference between $Y_1$ and $Y_0$.

The propensity score includes race/ethnicity, a dummy for Philadelphia to capture the differential incarceration rate between Philadelphia and Phoenix (Mulvey et al., 2007), year dummies, family background characteristics, proxies for previous criminal activity, outcomes before the treatment, risk perceptions, and a substance and alcohol abuse dummy. The probability of being incarcerated, conditional on committing a crime, is the product of three different probabilities: (i) the probability of being caught, (ii) the probability of being sentenced to incarceration conditional on being caught, and (iii) the probability of incarceration conditional on being sentenced. Since I work with a selected sample of juvenile offenders, I need controls that capture information about the last two probabilities.

First, for the probability of being sentenced to incarceration conditional on being caught, besides race, age, and location, I included the seriousness of the offence that brought the respondent into the survey (i.e. initial referral).[14] I also include variables that affect the criminal propensity of the individual like a proxy for criminal capital (Lochner, 2004; Loughran et al., 2013), family background,[15] the individual's perceptions about the risk associated with criminal activity, and self-control (Gottfredson and Hirschi, 1990; Duckworth and Seligman, 2017; Heckman et al., 2006).[16]

PTD has information that allows me to include measures for these elements. In the baseline interview, participants reported the age of criminal onset which I use as a proxy for criminal capital. I also include variables for family background such as dummies for father and mother being high school graduates, coming from a complete family, and criminal records from the parents. My measure of self-control comes from temperance as measured by the Weinberger Adjustment Inventory (WAI), which is defined as the ability to regulate emotional and behavioural impulses. The importance of temperance in predicting antisocial behaviour has been confirmed in different studies (Cauffman et al., 2005) and it is discussed in detail in Chapters 3 and 4. I also use a measure of the perceived probability of being caught when committing a

[14] These measures could reflect how dangerous the offender is perceived by the judge, which is a key element to consider when he/she is determining the severity of the punishment. In the propensity score, I include dummies for different types of crime in the initial referral. For example, Bhuller et al. (2016) include the type of crime in their model for the probability of incarceration.

[15] The literature that relates family background and crime is abundant. See for example Currie and Tekin (2012); Eriksson et al. (2016); Rowe and Farrington (1997); Pezzin (2004); Case and Katz (1991).

[16] An important predictor of criminal activity is labour market conditions (Lochner, 2004; Grogger, 1998; Imai and Krishna, 2004). However, given my definition of the treatment, in which individuals are in school years and have limited work experience, labour market conditions are not strong predictors for criminal behaviour.

crime. There is documented evidence about the relationship between risk perceptions and the propensity of criminal engagement (Lochner, 2007; Anwar and Loughran, 2011). In this way, perceptions are good predictors of crime which in turn is a good predictor of incarceration.

Altogether, these variables capture key elements considered by judges to determine whether an offender should be placed in a residential facility. These variables are part of the current algorithms used in the court system to determine the offender's threat to society and reoffending probability.[17]

Second, for the probability of incarceration conditional on being sentenced, PTD does not have direct proxies that capture such information. However, given the extensive set of controls provided by PTD, I expect that some of the previous covariates capture part of that variation, like race, criminal capital, and seriousness of the initial referral.[18]

Finally, since mental health (Frank and McGuire, 2010; Coker et al., 2014; Schubert et al., 2011; Hoeve et al., 2013; Lamb and Weinberger, 1998) and an individual's history of drug and alcohol abuse[19] are strong predictors of criminal engagement, preexisting differences in those variables across individuals may explain differences in incarceration probabilities. By including these elements in the propensity score, I ensure the distribution of observable characteristics is balanced across the control and treatment group.

Table 2.1 compares both groups in terms of the mean and standard deviation of the variables included in the propensity score. The last column of Table 2.1 presents the p-values associated with testing if the means are different between groups.

The differences are given by the proportion of males going to detention, the number of years of previous criminal experience, the pre-incarceration measures of depression, anxiety, somatization, and risk perception, and the proportion of people with substance abuse problems. Such pre-existing differences in characteristics generate a different distribution of propensity scores between the two groups. In particular, the mass of the distribution for the treatment group is concentrated at higher values, while it is concentrated at lower values for the control group. As a consequence, for high and low values of the propensity scores, only a few observations can be found in the other group. Therefore, some of my estimates for the effect of incarceration on mental health are noisy. Section 2.5 discusses the effects of incarceration calculated after I restrict the support to the region where both groups have a substantial number of observations.

To guarantee the *common support* assumption, I drop observations with a propensity score below the minimum observed in the treatment group and above the maximum observed in the control group. Table A.2 presents the estimates of the propensity score for this subsample.[20]

[17] See for example Kehl and Kessler (2017) and Knox (2017).

[18] It is important to clarify that in my data I observe the sentence outcome and the residential placement conditional on the sentence. Therefore, I do not need an exclusion restriction in my propensity score to separate the probability of being sentenced to incarceration conditional on being caught and the probability of incarceration conditional on being sentenced.

[19] In addition, the literature has found a strong correlation between mental health problems, drug and alcohol abuse, and crime (Saffer and Dave, 2002; Fridell et al., 2008; Schubert et al., 2011; Rajkumar and French, 1997).

[20] In the second column of Table A.2, I also included the predicted propensity score ($\hat{P}(X)$) among the regressors. If the quality of the matching is good, after controlling for $\hat{P}(X)$, the observable characteristics should not generate additional information regarding the probability of getting into incarceration. In fact, in Table A.2 none of the coefficients are statistically different from zero after controlling for the propensity score.

Table 2.1: Descriptive statistics: Means and standards deviation from control and treatment groups.

| Variables | Control | | Treatment | | p-value |
|---|---|---|---|---|---|
| Race | | | | | |
| White | 0.286 | (0.453) | 0.225 | (0.419) | 0.232 |
| Black | 0.354 | (0.48) | 0.437 | (0.498) | 0.143 |
| Hispanic | 0.298 | (0.459) | 0.275 | (0.448) | 0.653 |
| Other | 0.062 | (0.242) | 0.063 | (0.245) | 0.964 |
| Philadelphia | 0.484 | (0.501) | 0.542 | (0.5) | 0.317 |
| Male | 0.733 | (0.444) | 0.901 | (0.299) | 0.000 |
| High School Diploma | | | | | |
| Mother | 0.596 | (0.492) | 0.563 | (0.498) | 0.564 |
| Father | 0.522 | (0.501) | 0.408 | (0.493) | 0.049 |
| Criminal Background | | | | | |
| Mother | 0.143 | (0.351) | 0.155 | (0.363) | 0.769 |
| Father | 0.292 | (0.456) | 0.324 | (0.47) | 0.548 |
| Complete Family | 0.224 | (0.418) | 0.141 | (0.349) | 0.064 |
| Cognitive Ability | 87.422 | (13.858) | 86.690 | (13.039) | 0.637 |
| Prior Detention | … | … | … | … | … |
| Prior Criminal Experience | 0.434 | (0.993) | 0.712 | (1.115) | 0.023 |
| Initial Referral | | | | | |
| Person | 0.429 | (0.496) | 0.408 | (0.493) | 0.724 |
| Property | 0.248 | (0.433) | 0.218 | (0.415) | 0.538 |
| Drug | 0.161 | (0.369) | 0.148 | (0.356) | 0.745 |
| Other | 0.161 | (0.369) | 0.225 | (0.419) | 0.160 |
| Initial Disposition | | | | | |
| Residential Placement | 0.118 | (0.324) | 0.620 | (0.487) | 0.000 |
| Probation | 0.839 | (0.369) | 0.338 | (0.475) | 0.000 |
| Other | 0.025 | (0.156) | 0.035 | (0.185) | 0.597 |
| Baseline Interview | | | | | |
| Depression (std) | 0.120 | (1.113) | 0.440 | (1.388) | 0.027 |
| Anxiety (std) | 0.205 | (1.209) | 0.482 | (1.54) | 0.081 |
| Somatization (std) | 0.175 | (1.151) | 0.531 | (1.514) | 0.021 |
| Hostility (std) | 0.162 | (1.104) | 0.351 | (1.164) | 0.148 |
| Risk Perception | 59.017 | (26.283) | 48.913 | (28.896) | 0.002 |
| Temperance (std) | -0.168 | (0.972) | -0.298 | (0.918) | 0.233 |
| Substance Abuse | 0.478 | (0.501) | 0.641 | (0.481) | 0.004 |
| People | 161 | | 142 | | |

Standard deviation in parentheses. P-value is calculated for the alternative hypothesis.

## 2.4 Matching Estimates of the Impact of Incarceration on Mental Health

The Propensity Score Matching (PSM) estimates for ATE of being incarcerated appear in Table 2.2. As I mentioned before, I use different matching algorithms: (i) nearest neighbour matching with 5 neighbours, (ii) IPW, (iii) IPW with regression adjustment on the pre-incarceration level of the outcome (IPWRA), and (iv) MDID.[21]

[21] Appendix A contains the figures comparing the pre-incarceration tendency between the control and treatment group for all dimensions of mental health. Also, ATT estimates are in Tables A.3 and A.4. Given that most coefficients are not significant at conventional levels, I do not discuss these results.

Table 2.2: Propensity score matching estimates of ATE of incarceration: Full subsample.

| Matching algorithm | GSI | Depression | Anxiety | Somatization | Hostility |
|---|---|---|---|---|---|
| Propensity Score Matching | | | | | |
| 5 Nearest-neighbors | 0.116 | 0.295*** | 0.1520 | -0.0240 | 0.245*** |
| | (0.122) | (0.101) | (0.106) | (0.117) | (0.092) |
| IPW | 0.125 | 0.261** | 0.0540 | -0.1210 | 0.196** |
| | (0.111) | (0.106) | (0.109) | (0.111) | (0.099) |
| IPW with RA | 0.098 | 0.223** | 0.0260 | -0.1320 | 0.186* |
| | (0.117) | (0.108) | (0.108) | (0.111) | (0.097) |
| Observations | | | 303 | | |
| Control | | | 161 | | |
| Treatment | | | 142 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.3: MDID estimates of ATE of incarceration: Full subsample.

| Matching algorithm | GSI | Depression | Anxiety | Somatization | Hostility |
|---|---|---|---|---|---|
| Propensity Score Matching | | | | | |
| 5 Nearest-neighbors | 0.135 | 0.249** | 0.0390 | -0.0230 | 0.245*** |
| | (0.120) | (0.108) | (0.127) | (0.117) | (0.093) |
| IPW | 0.074 | 0.181* | 0.0090 | -0.1350 | 0.174* |
| | (0.113) | (0.108) | (0.104) | (0.109) | (0.1) |
| IPW with RA | … | … | … | … | … |
| | … | … | … | … | … |
| Observations | | | 303 | | |
| Control | | | 161 | | |
| Treatment | | | 142 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regarding depression, my PSM estimates range from 0.22 to 0.29 standard deviations and almost all of them are statistically significant at the 5% level. Therefore, a random person in my population suffers an increase of at least 0.22 standard deviations in the depression level due to the experience of incarceration. Regarding hostility, my estimates range from 0.18 to 0.24 standard deviations and most of them are also significant at the 5% level.

For anxiety and somatization, the effects are smaller. They range from 0.02 to 0.15 standard deviations in the case of anxiety and from -0.13 to -0.02 for somatization. This suggests modest improvements in the somatization dimension of mental health. However, the effects are imprecisely estimated and none of them are significant at conventional levels.

The ATE is also estimated using MDID (Heckman et al., 1997) and the results are in Table 2.3. As I mentioned in Section 2.3.1, this method exploits the longitudinal nature of the PTD data and uses difference-in-differences to control for fixed unobservable determinants of both incarceration and mental health. For depression, the MDID estimates are smaller than the PSM estimates. By assuming that fixed unobservable characteristics enter the outcome equation in an additively separate way, it is possible to conclude that fixed unobservable characteristics help to increase the effects of incarceration.

In the case of hostility, MDID and PSM produce almost the same interval of values. The similarity between the estimates suggests that, in the full support, fixed unobserved characteristics that affect both, incarceration and hostility, are properly controlled for by the set of covariates included in the propensity score. I discuss this issue further in Section 2.5.

Finally, note that for GSI the estimates are smaller for depression and hostility. The fact that estimates for anxiety are close to zero and for somatization are negative explains why GSI estimates are smaller. However, none of the estimates are significant at conventional levels and, therefore, it is not possible to derive further conclusions.

## 2.5 Are the Effects Different in the "Thick" Support?

In this section, I present the matching results in the "thick" support. Some estimates from the previous section have large standard errors. This is because the support condition only weakly holds for people with a low and high probability of incarceration. A small number of controls with high propensity scores provide the counterfactual for a much larger number of individuals with high propensity scores going into detention. The same is true for individuals with low propensity scores in the treatment group.

The mean propensity score for the treatment group is 0.57 while the control group has a mean of 0.37. The common standard deviation is 0.19. The distribution for the treatment group has more mass at higher values of the propensity score. For example, 33% of the treatment group lies above the 95th percentile of the control group and 22% of the control group lies below the 5th percentile of the treatment group. The common support gets thinner for higher and lower values of the propensity score.[22] Such a problem is corrected on the "thick" support where there is a sufficient number of observations for both the treatment and control groups (Black and Smith, 2004).

To examine how the "thin" support affects the estimates, I present estimates for the "thick" support region, which I define as the propensity scores $\hat{P}(X) \in [0.20, 0.65]$. In this region, there are at least 60% of the observations for both groups. Table 2.4 contains the estimates for PSM and Table 2.5 for MDID. There are some important lessons from this exercise. First, estimates for depression and hostility gain in terms of significance. The coefficients are now significant at the 5% level. In the case of GSI, the coefficient from IPW combined with difference-in-differences is also significant at the 5% level.

Second, the estimates for anxiety and somatization are not significant at conventional levels even in the "thick" support. I conclude that incarceration has an effect on only two dimensions of mental health: depression and hostility.

Third, while matching estimates for depression are similar to the ones from the complete subsample, the MDID estimates show a larger effect of incarceration that goes between 0.31 and 0.34 standard deviations.[23] This suggests that fixed unobservable characteristics attenuate

[22]To illustrate this point even further, consider that 3.5% of treated observations are below the 10th percentile of the distribution of propensity scores. Also, 3.7% of control observations are above the 90th percentile.

[23]These matching estimates are statistically different from the ones in the full support.

Table 2.4: Propensity score matching estimates of ATE of incarceration: “Thick” support.

| Matching algorithm | GSI | Depression | Anxiety | Somatization | Hostility |
|---|---|---|---|---|---|
| Propensity Score Matching | | | | | |
| 5 Nearest-neighbors | 0.079 | 0.22** | -0.0230 | -0.0940 | 0.201** |
| | (0.372) | (0.09) | (0.107) | (0.1) | (0.099) |
| IPW | 0.131 | 0.239** | 0.0000 | -0.0940 | 0.268** |
| | (0.100) | (0.11) | (0.109) | (0.099) | (0.107) |
| IPW with RA | 0.166 | 0.294** | 0.0250 | -0.0670 | 0.287*** |
| | (0.108) | (0.117) | (0.114) | (0.107) | (0.106) |
| Observations | | | 191 | | |
| Control | | | 107 | | |
| Treatment | | | 84 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.5: MDID estimates of ATE of incarceration: “Thick” support.

| Matching algorithm | GSI | Depression | Anxiety | Somatization | Hostility |
|---|---|---|---|---|---|
| Propensity Score Matching | | | | | |
| 5 Nearest-neighbors | 0.123 | 0.312** | 0.0090 | -0.0220 | 0.204** |
| | (0.171) | (0.13) | (0.071) | (0.138) | (0.08) |
| IPW | 0.205** | 0.345*** | 0.0760 | -0.0100 | 0.316*** |
| | (0.102) | (0.112) | (0.107) | (0.104) | (0.113) |
| IPW with RA | … | … | … | … | … |
| | … | … | … | … | … |
| Observations | | | 191 | | |
| Control | | | 107 | | |
| Treatment | | | 84 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

the effects of incarceration on depression in the “thick” support. As a consequence, the role of fixed unobservable characteristics is different in the full versus the “thick” support.

In principle, bigger estimates in the “thick” support can be the result of two different sources. First, there may be heterogeneous treatment effects with bigger impacts on the middle values of the distribution of propensity scores. Second, it may be that lingering selection on unobservables has a smaller effect for values of $\hat{P}(X)$ outside the “thick” support. However, given that my estimates are still bigger for depression after controlling for fixed unobservable characteristics, the existence of lingering selection does not seem to drive the results.

Fourth, for hostility, some of the coefficients for the “thick” support are bigger than the estimates for the full support. This conclusion remains after using MDID. However, the evidence for bigger effects for middle values of the incarceration probability is not as robust as for depression. The same conclusion holds for the GSI. Note that estimates for the GSI are smaller than for depression and hostility due to the influence of the other two dimensions of mental health: anxiety and somatization.

Altogether, these findings are compatible with the existence of heterogeneous treatment effects for depression with a bigger impact on the "thick" support. In Section 2.7, I discuss how ATE estimates vary with the probability of incarceration and show that bigger effects for the middle values of the propensity score are the result of not using people with the highest incarceration probability to estimate the effects.

## 2.6 Do the Effects Differ by Race and Gender?

In this section, I discuss how the estimates from the "thick" support change across different groups. One of the problems of using the "thick" support to estimate the effects by groups is the reduction in the sample size. Even though I have enough comparable units in both the treatment and control groups, the sample is not large enough to identify the effect for different race groups, facility types, or locations. I restrict the attention to the groups where I find precise effects: blacks and males.

Table 2.6 contains the results for blacks and Table 2.7 for males. First, in the "thick" support, incarceration increases depression more for a random black person than for a random person. My estimates are almost twice as big for blacks. Regarding males, my estimates are smaller. Almost all of them are around 0.22 standard deviations. For depression, the results suggest the existence of heterogeneous effects of incarceration with bigger impacts for blacks and smaller effects for males.

Table 2.6: Propensity score matching and MDID estimates of ATE of incarceration for blacks.

| | GSI | | Depression | | Hostility | |
|---|---|---|---|---|---|---|
| Matching algorithm | PSM | MDID | PSM | MDID | PSM | MDID |
| Propensity Score Matching | | | | | | |
| 5 Nearest-neighbors | 0.409* | 0.392*** | 0.635*** | 0.756*** | 0.391 | 0.432** |
| | (0.226) | (0.124) | (0.196) | (0.163) | (0.388) | (0.220) |
| IPW | 0.363** | 0.321 | 0.664 | 0.642** | 0.371 | 0.468** |
| | (0.177) | (0.316) | (0.537) | (0.279) | (0.267) | (0.213) |
| IPW with RA | 0.334** | … | 0.650*** | … | 0.441** | … |
| | (0.157) | … | (0.128) | … | (0.212) | … |
| Observations | | | 71 | | | |
| Control | | | 39 | | | |
| Treatment | | | 32 | | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

My estimates also show that incarceration increases hostility more for blacks and the effects for males are at least equal to the ones in the "thick" support. The same conclusions hold for GSI. Together, the results indicate that incarceration is particularly detrimental to the mental health of blacks.

Table 2.7: Propensity score matching and MDID estimates of ATE of incarceration for males.

| | GSI | | Depression | | Hostility | |
|---|---|---|---|---|---|---|
| Matching algorithm | PSM | MDID | PSM | MDID | PSM | MDID |
| Propensity Score Matching | | | | | | |
| 5 Nearest-neighbors | 0.200** | 0.186** | 0.227** | 0.187** | 0.271*** | 0.263*** |
| | (0.088) | (0.075) | (0.101) | (0.086) | (0.097) | (0.094) |
| IPW | 0.174* | 0.203** | 0.212* | 0.236** | 0.311*** | 0.316*** |
| | (0.097) | (0.091) | (0.109) | (0.106) | (0.109) | (0.115) |
| IPW with RA | 0.192** | … | 0.222** | … | 0.311*** | … |
| | (0.089) | … | (0.104) | … | (0.108) | … |
| Observations | | | 165 | | | |
| Control | | | 91 | | | |
| Treatment | | | 74 | | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

## 2.7 Do the Effects Vary with the Incarceration Probability?

One common approach to analyze the existence of heterogeneous treatment effects is to check the ATE varies with participation probabilities (Lechner, 2002; Morgan, 2001; Xie and Wu, 2005). I analyze such variation in two ways. First, I rank the subsample for estimation according to the propensity score and estimate the effects of incarceration in each tercile of the distribution. A common practice in the literature is to split the sample into five or more groups (Morgan, 2001; Brand and Xie, 2010; Xie and Wu, 2005; Stürmer et al., 2014) and compare the estimates from each group. However, given the size of my sample, doing so would leave only a few observations in each stratum to precisely estimate the effects.

Second, I use kernel-smoothed regressions of different outcomes on incarceration probabilities for people in the control group versus people in the treatment group (Lechner, 2002; Xie et al., 2012). I present only the results for depression, hostility, and GSI because I did not find conclusive evidence for anxiety and somatization. It is important to emphasize that the following results should be taken with caution in regards to the regions of the propensity score where there is not much overlap between the control and the treatment group, given the discussion from Section 2.5. In particular, the heterogeneity I find in the first and third terciles may be influenced by the small number of units used to construct the counterfactual. As a result, my estimates in those regions may not be as reliable as the ones in the second tercile.

Table 2.8 presents the MDID estimates for GSI, depression, and hostility in each tercile. For depression and GSI the effects of incarceration decrease by tercile. People with the lowest probability of incarceration suffer the worst consequences. According to my estimates for depression, the effects in the first tercile are more than two times bigger than effects for the full support. Also, the effects in the second tercile are comparable to the ones from the "thick" support. Finally, people in the third tercile have a decrease in the measure of depression that is almost equivalent to the increase in the first tercile. The bigger effects in the "thick" support come from eliminating the upper tail of the distribution. In my sample, people who are more likely to go to detention are the ones with more criminal experience, worse mental health, and higher prevalence of drug and alcohol abuse problems (Table A.5). Interestingly, their mental health improves after spending time in a residential facility.

Table 2.8: MDID estimates of ATE of incarceration (by terciles).

| | GSI | | | Depression | | | Hostility | | |
|---|---|---|---|---|---|---|---|---|---|
| Matching Algorithm | 1st | 2nd | 3rd | 1st | 2nd | 3rd | 1st | 2nd | 3rd |
| Propensity Score Matching | | | | | | | | | |
| 5 Nearest-neighbors | 0.434** | 0.171 | -0.684*** | 0.669*** | 0.279 | -0.52** | 0.328 | 0.172* | 0.034 |
| | (0.197) | (0.155) | (0.233) | (0.218) | (0.194) | (0.221) | (0.242) | (0.101) | (0.191) |
| IPW | 0.435** | 0.284* | -0.807** | 0.573** | 0.356*** | -0.649** | 0.278 | 0.525* | 0.037 |
| | (0.179) | (0.16) | (0.37) | (0.26) | (0.127) | (0.271) | (0.4) | (0.314) | (0.162) |
| IPW with RA | 0.387** | 0.213 | -0.761 | 0.557*** | 0.287** | -0.536* | 0.221 | 0.439** | 0.048 |
| | (0.189) | (0.146) | (0.481) | (0.162) | (0.131) | (0.298) | (0.348) | (0.219) | (0.155) |
| Observations | 101 | 101 | 101 | 101 | 101 | 101 | 101 | 101 | 101 |
| Control | 81 | 61 | 27 | 81 | 61 | 27 | 81 | 61 | 27 |
| Treatment | 20 | 40 | 74 | 20 | 40 | 74 | 20 | 40 | 74 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

One potential explanation for such an improvement is that incarceration reduces the probability of substance abuse.[24] To the extent that depression is strongly associated with substance abuse (Calcaterra et al., 2014), a reduction in the probability of abusing translates into improvements in mental health. Another potential explanation is that PTD participants with significant mental health problems are more likely to receive medical help and additional psychological services while in detention, especially in the presence of substance use problems (Mulvey et al., 2007).

Regarding the GSI, the results are similar to the ones from depression. Incarceration causes: (i) deterioration in the overall mental health for people in the first two terciles with a higher impact in the first tercile and (ii) an improvement in the overall mental health for people with the highest probability of going into detention. The effects for the first two terciles are smaller than for depression, but the improvement in the third tercile is bigger. Finally, there is an important gain in terms of significance for the GSI and depression estimates.

For hostility, my estimates show that everybody deteriorates in this dimension of mental health and the effects in the third tercile are smaller than in the other two. However, my estimates for the first and third tercile are not significant at conventional levels. In addition, estimates from IPW and IPWRA are bigger for the middle values of the propensity score than for the full support.

Figure 2.1 shows kernel-smoothed regressions of the incarceration probabilities on GSI, depression, and hostility for people in the control group versus people in the treatment group.[25] The regressions are estimated using a Gaussian kernel and the rule-of-thumb bandwidth. For depression and GSI, the results in Figure 2.1a and Figure 2.1b, respectively, confirm the conclusions from estimating ATE by terciles. Incarceration causes an increase in depression and GSI for people with a low probability of incarceration and a decrease in the same measures for

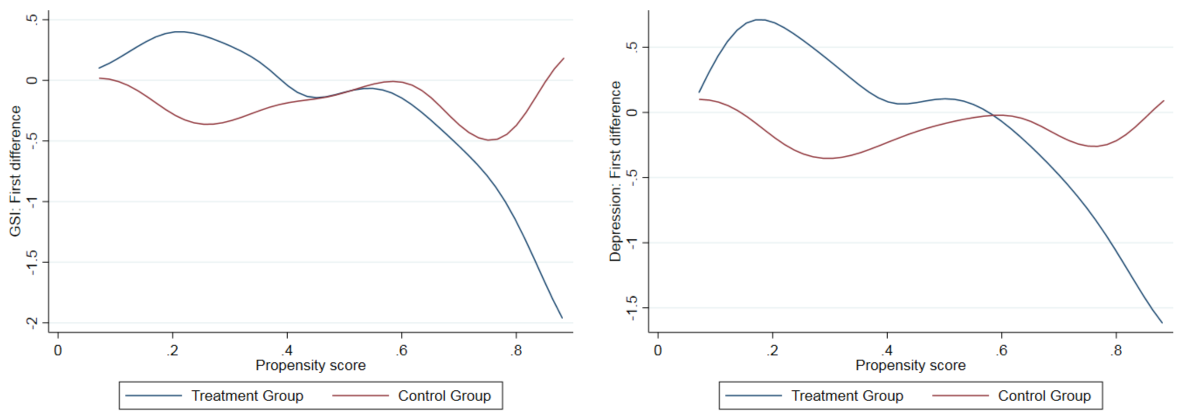
[24] In my subsample, a person from the treatment group is 17.8 percentage points less likely to report a substance abuse episode at age 18 than a person from the control group. I find this difference after controlling for several demographic characteristics and the history of substance abuse. The difference is significant at the 5% level.

[25] The dependent variable is the difference between the after the treatment level and the before the treatment level.

people with a high probability of incarceration. For both dimensions, the effects are bigger for propensity score values smaller than 0.35 (which are roughly the values from the first tercile). Also, the higher the probability of incarceration, the bigger the improvement for depression and the overall mental health.

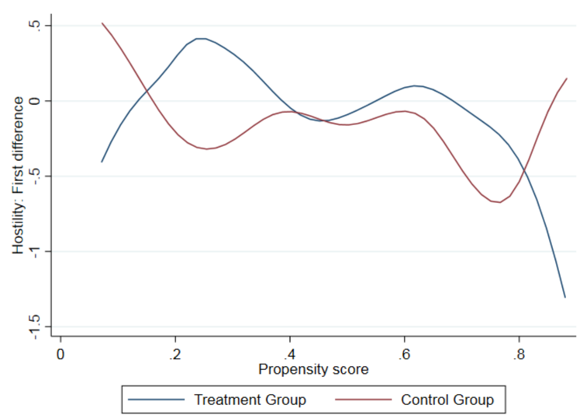
Figure 2.2a shows that incarceration causes an increase in hostility for most of the interval as was discussed before. Also, it shows the effect is bigger for values between 0.2 and 0.4.

Figure 2.1: Nonparametric regression of the incarceration probabilities on mental health in respective subsamples.




(a) GSI (b) Depression

Figure 2.2: Nonparametric regression of the incarceration probabilities on mental health in respective subsamples.




(a) Hostility

My results suggest that standard estimates, that do not recognize the existence of heterogeneous effects, understate the effects of incarceration on mental health for people with low a incarceration probability. Since they are the ones suffering the worst consequences of incarceration, cost-effective policy design should pay special attention to this group.

Also, people with more criminal experience, substance abuse problems, and bad mental health may benefit from spending time in an institutionalized environment with regard to mental health. As I argue in Chapters 3 and 4, incarceration produces several negative effects including the creation of criminal capital, the deterioration of self-control, and the depreciation of human capital.

## 2.8 Effects of incarceration over subsequent criminal participation: A back of the envelope exercise

In this section, I present a back-of-the-envelope exercise to illustrate the effect of changes in depression and hostility, produced by incarceration, on criminal engagement. In particular, I answer the following questions: (i) What would the associated increase in crime propensity be for an average person in the control group, between 18 and 19 years old, if he/she had experienced a change in depression and hostility that is equivalent to the one experienced by someone who went into incarceration? (ii) What would the crime propensity be for an average person in the treatment group, between 18 and 19 years old, if he/she had not committed the crime that led to incarceration?

To answer both questions, I only use people in the control group because they do not have any incarceration experience by age 18 and so changes in crime propensity are not associated with the experience of incarceration.[26] To deal with the potential simultaneity between mental health and crime, I use lagged (by one period) measures of depression and hostility. By further assuming that depression and hostility are exogenous to crime propensity and CIA extends to the criminal participation that happens between 18 and 19 years old, it is possible to isolate the effect of mental health.

I estimate probit models for criminal participation using the same set of covariates from the propensity score.[27] To answer the second question, I also need to account for the fact that controls differ in multiple observable dimensions from the treatment group. To mitigate this fact, I estimate the probits using the IPW from the propensity score, which makes both groups comparable.

Table A.6 presents the probit estimates and Table 2.9 shows the average marginal effects of hostility and depression on criminal participation.[28] From the estimates, an increase of one standard deviation in depression is associated with an increase of 9 to 18 percentage points in the probability of committing a crime, depending on the model. All of the effects are significant at the 10% level. Regarding hostility, an increase of 1 standard deviation is associated with an increase of 7 to 12 percentage points in the probability of committing a crime but the effects are not significant at conventional levels when the probit is estimated using IPW weights.

[26] For example, there is documented evidence showing that incarceration creates criminal capital (Bayer et al., 2009; Nguyen et al., 2017), discrimination in the labour market (Pager, 2003), and influences the search frictions for crime and employment (Mancino, 2018; Imai and Krishna, 2004).

[27] I estimate two specifications. The basic one has the same covariates as the propensity score. The second one has a more complete specification that also includes accumulated semesters of schooling, legal experience, criminal experience, year dummies, and earnings from illegal activities in the previous period.

[28] The predicted probabilities used to calculate the marginal effects are calculated using the observed and mean values of the covariates.

Table 2.9: Estimated marginal effect without IPW weights

| | No Weights | | | |
|---|---|---|---|---|
| Variables | Depression | | Hostility | |
| Depression | | | | |
| At Observed Values | 0.096*** | 0.093*** | 0.088*** | 0.089*** |
| | (0.035) | (0.036) | (0.032) | (0.032) |
| At Means | 0.134*** | 0.13** | 0.123*** | 0.124*** |
| | (0.051) | (0.051) | (0.046) | (0.047) |
| Monetary Benefits | No | Yes | No | Yes |
| Observations | 321 | | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.10: Estimated marginal effect with IPW weights

| | IPW Weights | | | |
|---|---|---|---|---|
| Variables | Depression | | Hostility | |
| Depression | | | | |
| At Observed Values | 0.124** | 0.122** | 0.0810 | 0.0770 |
| | (0.061) | (0.061) | (0.06) | (0.061) |
| At Means | 0.18* | 0.178* | 0.1180 | 0.1130 |
| | (0.092) | (0.092) | (0.089) | (0.09) |
| Monetary Benefits | No | Yes | No | Yes |
| Observations | 321 | | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

To answer the first question, note that incarceration increases depression by at least 0.18 standard deviations for the full subsample and at least 0.22 standard deviations in the "thick" support. Therefore, an average person in the control group would have an increase of at least 1.62 percentage points in the probability of committing a crime if he/she had experienced a change in depression that is equivalent to the one experienced by someone who went into incarceration.[29] Regarding hostility, my estimates show that incarceration increases hostility by at least 0.17 standard deviations in the full support and 0.2 standard deviations in the "thick" support. Therefore, an average person in the control group would have an increase of at least 1.36 percentage points in the probability of committing a crime if he/she had experienced a change in hostility that is equivalent to the one experienced by someone who went into incarceration. If I further assume the two effects are additively separable, the equivalent deterioration of mental health is associated with an increase of at least 2.98 percentage points in the probability of criminal participation.

To answer the second question, I use the estimates from the probit with IPW weights. Regarding depression, an average person in the treatment group would have an increase of at least 2.16 percentage points in the probability of committing a crime if he/she had not committed the crime that led to incarceration. For hostility, that increase would be 1.19 percentage points. If I assume that the two effects are additively separable, the deterioration of mental health is associated with an increase of at least 3.35 percentage points in the probability of criminal participation.

Regarding this exercise, I would like to emphasize a few things. First, previous assumptions made to illustrate how changes in mental health affect subsequent criminal participation are rather restrictive. Second, my results are silent regarding the influence of mental health problems over schooling decisions and labour market outcomes.[30] To the extent that these choices affect the opportunity cost of crime (Grogger, 1998; Lochner and Moretti, 2004; Lochner, 2004), changes in mental health may also impact criminal engagement through schooling and employment. In Chapter 4, I estimate a dynamic model of employment, schooling, and crime where decisions are affected, among other things, by mental health, and mental health is affected by previous choices and the experience of incarceration. With this model in hand, I illustrate the effects of changes in mental health on subsequent criminal activity without invoking some of the restrictive assumptions used here. Also, I analyze changes in criminal behaviour over the life-cycle and not only for a particular age. My results from Chapter 4 suggest that mental health is not a key driving mechanism of long-run criminal participation once other forces, in particular, criminal capital, are taken into account.

## 2.9 Conclusion

In this chapter, I investigate the link between incarceration and mental health for disadvantaged young individuals. I take advantage of the rich data provided by the PTD survey and estimate,

[29] This the number is calculated by multiplying the lowest marginal effect times the lowest effect from my matching estimates. In this particular case, $0.0162 = 0.18 \cdot 0.09$.

[30] There is important literature linking mental health problems, employment and human capital accumulation (Reynolds et al., 2007; Ettner et al., 1997; Cook et al., 2009; Peng et al., 2016).

using different propensity score matching algorithms, the causal effect of incarceration on several dimensions of mental health.

According to my results, incarceration has important consequences on the deterioration of two dimensions of mental health: depression and hostility. Such effects are bigger for people with a low incarceration probability. Also, my estimates show that people with the highest incarceration probability experience a reduction in depression and GSI. I find bigger impacts on blacks and smaller effects for males in the case of depression. Finally, I show evidence that the deterioration of mental health, equivalent to that caused by incarceration, is related to an increase in crime propensity between ages 18 and 19.

My results provide evidence about the existence of a negative by-product of incarceration that feedbacks into reoffending. As a consequence, crime-fighting strategies that combine law enforcement and health policies have the potential of mitigating some of the negative consequences of incarceration. Furthermore, given the existence of heterogeneous effects, interventions that target the groups suffering the worst consequences of incarceration may be cost-effective mechanisms to reduce reoffending and mental health problems.

# Chapter 3

# Personal Capabilities, Human Capital Accumulation and Criminal Behavior

## 3.1 Introduction

A robust finding in the literature studying criminal engagement over the life-cycle is the strong positive correlation between past and future offending. The issue of reoffending has alarming proportions in the U.S. According to the statistics from the Bureau of Justice Statistics, in 2002 around six-hundred thousand people were released from incarceration and roughly two-thirds of those released will be rearrested within three years (Langan and Levin, 2002). Because of this, crimes by former inmates alone account for a substantial share of current and future crime. Reoffending rates are also staggering for juveniles. Fagan (1996) reported rearresting rates above 40% in the two years after release while Seigle et al. (2014) claim that reoffending rates could be as high as 75% within three years of release.

Many different approaches have been taken to understand the continuity in criminal behaviour, but no definitive answer has been found yet.[1] Some researchers have found that part of the observed continuity in offending is because, at a very early age, individuals have personal capabilities (i.e. individual heterogeneity) that affect their propensity to engage in crime. Examples of such characteristics are mental health problems (Coker et al., 2014; Fridell et al., 2008), neuropsychological deficits (Moffitt, 1993), lower self-control (Gottfredson and Hirschi, 1990), and antisocial traits (Caspi et al., 1994).

There is also evidence showing that previous choices have the potential of transforming the offender's life circumstances in such a profound way that it alters the probability that subsequent criminal acts will occur (i.e. state dependence). Choices like getting married (Laub and Sampson, 1998), gang affiliation (Barnes et al., 2010), dropping out from school (Thornberry et al., 1985), and incarceration (Bayer et al., 2009; Mueller-Smith, 2015; Bhuller et al., 2016), have received considerable attention in the literature.

A more comprehensive approach in the literature has studied both individual heterogeneity and previous choices to understand the continuation of criminal behaviour (Nagin and Paternoster, 1994; Land and Nagin, 1996; Laub and Sampson, 1998). For example, Mancino et al. (2016) use PTD data to separately identified the role of state dependence, previous choices and

[1] See for example the excellent review by Nagin and Paternoster (2000).

individual heterogeneity. In this chapter, I follow this approach and document how previous choices and personal capabilities are related to crime. The purpose of this exercise is to provide stylized facts to guide the modelling choices for the dynamic model I develop and estimate in Chapter 4 to analyze the life-cycle choices of juvenile offenders.

To provide a more complete picture of the role of both elements, I enrich the analysis in two ways. First, given that employment and schooling choices affect the opportunity cost of crime (Grogger, 1998; Lochner and Moretti, 2004; Lochner, 2004), I document how these two elements are related to labour market outcomes and human capital accumulation decisions (e.g. high school graduation). Second, I investigate if previous choices affect the evolution of personal capabilities. In this way, I provide evidence on a dynamic dimension of individual heterogeneity under which personal capabilities both influence and are influenced by schooling decisions, labour market participation, and criminal trajectories. The findings from Chapter 2 can be seen as an example of such dynamic interplay.[2]

To document the dynamic interplay between personal capabilities and choices, I also use the PTD survey. PTD was specifically designed to study the evolution of criminal behaviour. The survey pays special attention to the choices made by juvenile offenders as they transition into adulthood. As I describe in Chapter 2, it has detailed information about incarceration spells, employment, school attendance and criminal engagement. It also has a rich set of measures about personal characteristics, family background and offending history.

For the analysis, I characterize personal capabilities with two major components: mental health and noncognitive skills. For mental health, I work with the same five measures from Chapter 2: GSI, depression, anxiety, somatization, and hostility as measured by the BSI test. Regarding noncognitive skills, I focus on self-control. There is strong evidence pointing to the existing relationship between self-control and labour market outcomes, schooling and crime (Fletcher, 2012; Cook et al., 2009; Duckworth and Seligman, 2017; Heckman et al., 2006).

For this chapter, my results come from four different empirical exercises. First, I estimate linear probability models with fixed effects for high school graduation, unemployment (i.e. home production), and criminal engagement on personal capabilities, other personal characteristics, and previous choices (i.e. sector-specific experience, education, and incarceration records). Second, I estimate Mincer equations with fixed effects for wages and illegal earnings on personal capabilities and previous choices. In this way, I investigate if there is state dependence on employment and crime through returns to experience and education. Third, I estimate linear models with fixed effects for personal capabilities on previous choices and the experience of incarceration. These results allow me to evaluate if personal capabilities change in response to previous choices. Fourth, I estimate linear regressions of the change in personal capabilities on changes in criminal experience, legal experience, and high school graduation.

[2]As I documented in Chapter 2, mental health problems are associated with criminal behaviour and incarceration during adolescence. But also, incarceration contributes to exacerbating mental health problems. There is evidence showing that people with mental health problems are less likely to graduate from high school (Fletcher, 2010). To the extent that high school dropouts end up earning lower wages, they are also more likely to re-offend (Aizer and Doyle, 2015).

The omission of such feedback effects has important implications for the analysis of crime-fighting strategies. First, analysis of law-enforcement policies (e.g. harsher punishment) may underestimate the effects over reoffending. Second, analysis of alternative crime-fighting strategies, like high school subsidies, training programs and wage incentives, may conclude that individuals with criminal background are less responsive to such programs which overestimate inequality in policy effects.

To deal with the endogeneity problem associated with estimating simultaneous equations, I use three strategies. First, as in Mancino et al. (2016), I include a rich set of individual-level characteristics related to labour market outcomes, schooling decisions, and crime. Many of these variables are not commonly available in datasets used to study the continuation of criminal behaviour and therefore would typically end up included in the error term.[3] Second, I estimate models with fixed effects. Given the longitudinal nature of the PTD survey, I can control for fixed unobservable characteristics affecting both the evolution of personal capabilities and choices. Third, I assume that lagged personal capabilities affect choices and that errors, once I control for fixed effects, are not serially correlated.[4] The idea is to disentangle whether the correlations observed in the data are structural and hence they should be part of the dynamic model in Chapter 4.

Based on the analysis, I reach three conclusions. First, there is a negative relationship between criminal engagement and mental health as was documented in Chapter 2.[5] I also find that self-control is not directly associated with criminal engagement as previous studies have suggested. However, self-control is an important element for human capital accumulation. Also, depression has a negative association with high school graduation.

Second, there is evidence consistent with the existence of sector-specific human capital. In particular, I find that legal experience and education have returns only in the legal sector, while criminal experience and incarceration records have returns only in the illegal sector.

Third, personal capabilities change over time as a consequence of previous choices and the experience of incarceration. Using a larger subsample than the one from Chapter 2, I confirm the positive association between incarceration, criminal experience and mental health problems. Regarding legal sector choices, high school graduation is associated with more self-control while employment is associated with better mental health. Finally, more criminal experience is associated with worse mental health in the medium-run. These findings inform the modelling choices adopted in Chapter 4.

The remainder of the chapter proceeds as follows. Section 3.2 adds additional details about the PTD survey beyond those provided in Section 2.2. It also presents the classification criteria I use to define detention status, occupational choice, and some descriptive statistics on how personal capabilities relate to occupational choice and educational level. These data are the same as used in Chapter 4 to estimate the dynamic model. Section 3.3 presents the estimates of linear probability models for criminal participation, unemployment,[6] and high school graduation. Section 3.4 discusses the existence of returns to sector-specific experience and personal capabilities in wages and illegal earnings. Section 3.5 presents evidence on how changes in personal capabilities are related to high school graduation, legal and criminal experience, and

[3]For example, NLSY97 is a dataset commonly used to study criminal behaviour (Merlo and Wolpin, 2015; Lochner, 2007). NLSY97 does not have detailed measures on illegal earnings, the intensive margin of criminal activity, and noncognitive and social skills. Another example is the Philadelphia Birth Cohort Study which has been used by Imai and Krishna (2004), Tauchen et al. (1994), among others. Even though this survey has extensive measures of criminal engagement, it also lacks information on mental health, noncognitive skills and illegal earnings.

[4]Mancino et al. (2016) also assume that errors are serially uncorrelated. However, they do not use models with fixed effects for the analysis.

[5]It is important to emphasize that conclusions in this chapter are not restricted to a particular age as they were in Chapter 2 and they also cover different types of crimes.

[6]I use the terms unemployment and home production interchangeably.

the experience of incarceration. Finally, Section 3.6 concludes.

## 3.2 Data

### 3.2.1 PTD Survey

For the analysis, the data come from the PTD study, which was described in Chapter 2. Besides the information previously provided, it is important to add that, in the baseline interview, several measures of cognitive skills were collected. From these measures, I use the full-scale measure of the Wechsler Abbreviated Scale of Intelligence (WASI) which produces an estimate of general intellectual ability based on two subtests: vocabulary and matrix reasoning.[7]

As I mentioned before, during the follow-up interviews, each participant completed monthly calendars covering the period between the current and the last interview, i.e. recall period. These calendars have monthly information about money earned from illegal activities,[8] wages, monthly residential placements, the facility type where those placements took place,[9] and school enrolment.

Altogether, this information allows me to construct a monthly history for each participant regarding detention status, the facility type where the respondent spent time (adult facility vs. juvenile facility), employment, enrolment, crime, wages and illegal earnings. To have a computationally feasible number of periods to estimate the model, I collapse such information by semesters and merge it with data about cognitive and noncognitive skills, and mental health. To minimize the number of missing observations, I use only individuals who completed all of the follow-up interviews. Finally, since this study focuses on adolescents, I do not include in my sample individuals who were over 18 years old when the first measures were collected. As a result, I ended up with a balanced panel of 779 individuals, all of them with 14 semesters of observations.

### 3.2.2 Definitions

I classify the data among different mutually exclusive and exhaustive alternatives based on the following procedure.

#### Detention Status

I first determine if the respondent was in detention or not during the semester. A respondent $i$ is in detention if the majority of months are spent in a detention facility. If the number of months in detention is equal to the months in the community, I also classified $i$ as being in detention.

[7]The WAIS and its abbreviated version, the WASI, have for the past several decades been the most commonly used IQ tests (Borghans et al., 2008).

[8]The exact question is: "You mentioned that you had made money during the past N months from ways besides working. Did you make any money during this month from activities that are illegal?". In case the answer was affirmative, additional questions were presented like the months when the antisocial activities took place, the type of crimes committed, the number of weeks worked, and the weekly money earned.

[9]The survey provided different levels of classification for residential facilities. The more general is a two-category classification: juvenile and adult. I adopt these categories in this study.

Then, I establish the facility type where $i$ spends most of the time. It was not unusual for respondents to be transferred between facilities during a detention spell. To deal with such cases, I classify $i$ as being in a juvenile facility if $i$ spends at least the same amount of time in the juvenile facility as in the adult facility. After the detention status is defined, I classify individuals according to their activity choice in each period.

### Activity Choice

The observations for each period were constructed using monthly data on employment, criminal engagement, and school attendance. For 80% of the observations, only one alternative was chosen during a particular semester, and therefore the individual was classified as taking that alternative for the whole semester. In the remaining 20% of the observations, individuals chose several alternatives in a particular semester. For those cases, any rule used to create biannual data on choices is somewhat arbitrary. I use the following hierarchical rules to classify choices into mutually exclusive alternatives.

**Not in Detention:** For those not considered in detention, there are four alternatives: work, crime, school, and home production. Home production is the residual category, meaning $i$ is classified as choosing home production if $i$ cannot be classified in the other three categories. To be classified as choosing work, crime, or school, the total number of months, during the semester, in that activity has to be greater than zero and:

1. **Work:** $i$ is classified as choosing to work if the total number of months working is strictly greater than the total number of months with criminal engagement and the total number of months in school or any vocational training.[10]

2. **Crime:** $i$ is classified as choosing crime if the total number of months with criminal engagement is greater than or equal to the total number of months in school or any vocational training and the total number of months working.

3. **School:** $i$ is classified as choosing school if the total number of months in school or any vocational training is greater than or equal to the total number of months working, and strictly greater than the total number of months committing a crime.

4. **Home Production:** $i$ is classified in home production if and only if, the total number of months in crime, school, and working is equal to zero.

Given the previous rules, if $i$ chooses crime for the same total number of months as school and work, $i$ is classified as choosing crime for the whole semester. If the total number of months in school is greater than the total number of months in crime and equal to the total number of months working, $i$ is classified as choosing school for the whole semester.

[10] In a given month, $i$ is considered as being employed if $i$ worked at least 20 hours per week on average.

**Detention:** There are two alternatives when $i$ is in detention: crime and training. To be classified as choosing crime, $i$ has to participate in criminal activities for at least one month and:

1. **Crime:** $i$ is classified in crime if the total number of months with criminal engagement (while in detention) is greater than or equal to the total number of months in school or any vocational training (while in detention) and the total number of months working (while in detention).[11]

2. **Training:** $i$ is classified in training if the total number of months in school or any vocational training (while in detention) or working (while in detention) is greater than the total number of months in crime (while in detention).[12] Also, $i$ is classified in training, if the total number of months with criminal engagement is equal to zero.

#### Wages and Criminal Earnings

The calendars specify the months in which the respondent was employed. For those months, there is also information regarding the average hourly wage and the number of worked hours per week. To calculate the potential monthly wage for the individual, I use only the observations (i.e. months) where $i$ works on average more than 20 hours per week. Then, I multiply the reported hourly wage by 160 (40 hours per week times 4 weeks in a month). Finally, the potential monthly wage for the semester is the average of the potential monthly wages.

Regarding illegal earnings, the calendars contain information about the average amount of money earned from illegal activities per week. To calculate the potential monthly equivalent, I multiply the average weekly earnings by 4. The potential monthly illegal earnings for the semester is the average of the illegal earnings at the monthly level.[13]

### 3.2.3 How personal capabilities relate to choices: Descriptive statistics

To characterize the relation between personal capabilities and choices, I first present some descriptive statistics. Figures 3.1 and 3.2 show the averages from each measure by schooling level and occupational choice.

From Figures 3.1 and 3.2, it is possible to derive the following conclusions. First, people without a high school diploma have worse mental health. The higher the educational level, the higher the self-control (Heckman et al., 2006; Duckworth and Seligman, 2017).

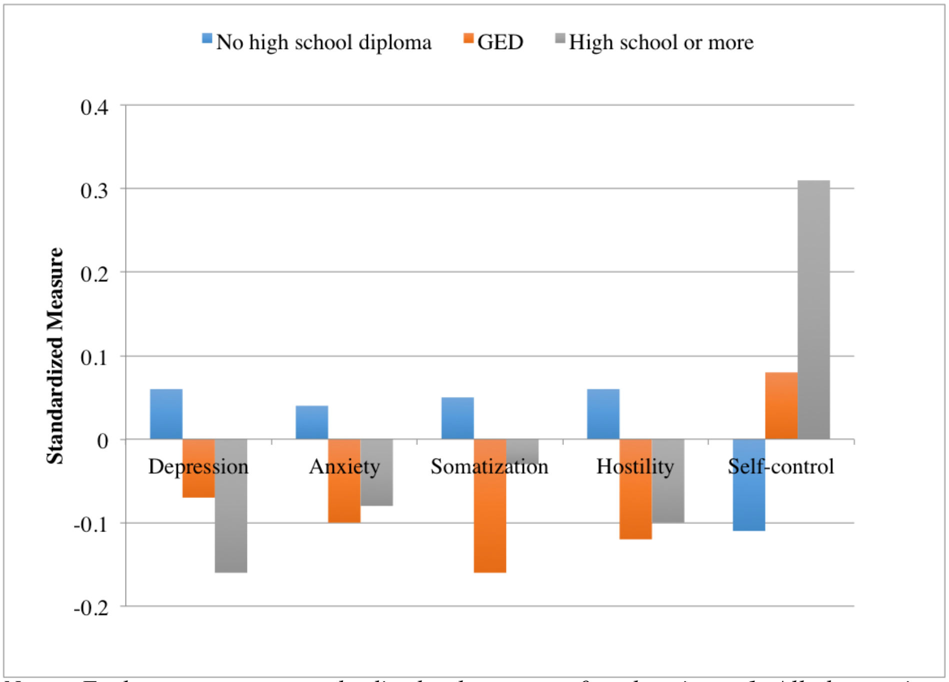
Second, it is clear that people in the criminal sector have worse mental health. Among criminals, self-control is lower (Gottfredson and Hirschi, 1990). In Section 3.3, I explore if the observed association between personal capabilities, schooling and crime remains after controlling for a rich set of variables and fixed unobserved characteristics.

[11] At the monthly level, only 2.3% of the observations that were classified as in detention, reported working in the legal sector.

[12] Even though employment is not a common choice while in detention, school is. In fact, at the monthly level, for 42% of the observations classified as "in detention" people reported receiving education or other forms of vocational training. Mulvey et al. (2007) documented people from PTD getting job skills training while in detention.

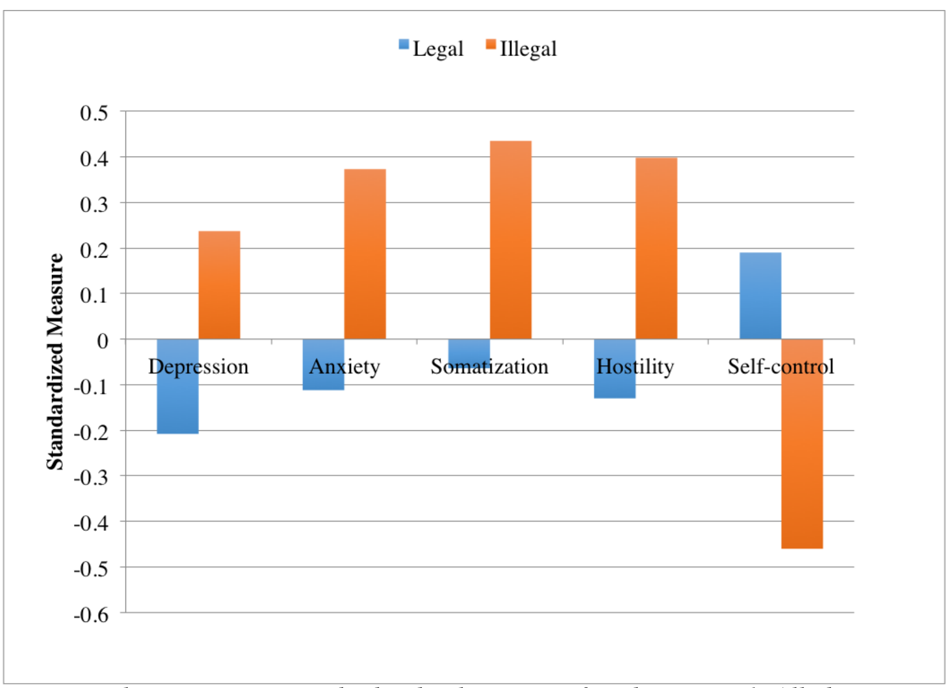
[13] At the monthly level, only 2.4% of the observations in which the subject reports earning money from illegal activities, earnings are missing or equal to zero.

Figure 3.1: Average personal capabilities by educational level.




*Notes: Each measure was standardized to have mean 0 and variance 1. All observations with valid scores on tests for mental health and self-control are used to calculate the average. Each subject is classified according to the highest educational level.*

Figure 3.2: Average personal capabilities by sector.




*Notes: Each measure was standardized to have mean 0 and variance 1. All observations with valid scores on tests for mental health and self-control are used to calculate the average. Sectors are mutually exclusive.*

## 3.3 Personal Capabilities and Choices

In this section, I analyze how personal capabilities are related to criminal engagement, high school graduation, and home production.[14] For criminal engagement, I consider (i) income crime and (ii) any type of crime. For both types of crime, I take a step back from the previous classification. Here, I define a dummy variable that takes the value of one if the respondent engages in any income-related crime or any crime during the semester, and zero otherwise. I restrict the attention to these choices because for employment and schooling I did not find conclusive evidence.

For each choice, I estimate the following linear model with fixed effects:

$$y_{it} = \boldsymbol{x}_{it1}\boldsymbol{\beta} + \boldsymbol{h}_{it-1}\gamma + c_i + u_{it}, \tag{3.1}$$

where $y_{it}$ is a dummy variable that represents the choice under analysis, $\boldsymbol{x}_{it1}$ are observed individual characteristics, $\boldsymbol{h}_{it-1}$ are personal capabilities, $c_i$ represents the fixed unobserved heterogeneity, and $u_{it}$ is the idiosyncratic error. In $\boldsymbol{x}_{it1}$, I include controls for substance abuse, age, age squared, incarceration so far, observed accumulated periods of education, legal experience and criminal experience. Periods of sector-specific experience and education are measured in semesters. Table 3.1 presents the results.

Table 3.1: Estimated parameters from linear probability models with fixed effects for crime, high school graduation and home production.

| Variable | Income crime | All crime | Home production | High school graduation |
|---|---|---|---|---|
| Mental health (previous period) | | | | |
| Depression | -0.001 | -0.004 | -0.027** | -0.012* |
| | (0.013) | (0.015) | (0.012) | (0.007) |
| Anxiety | -0.006 | -0.015 | -0.007 | 0.003 |
| | (0.013) | (0.015) | (0.012) | (0.008) |
| Somatization | 0.005 | 0.002 | 0.014 | -0.003 |
| | (0.012) | (0.013) | (0.011) | (0.007) |
| Hostility | 0.026** | 0.027** | -0.003 | 0.009 |
| | (0.012) | (0.014) | (0.011) | (0.007) |
| Self-control (previous period) | -0.017 | -0.009 | -0.018 | 0.015** |
| | (0.013) | (0.015) | (0.012) | (0.008) |
| Observations | 2,859 | 2,859 | 2,859 | 2,431 |
| R-squared | 0.024 | 0.062 | 0.055 | 0.171 |

Notes: Standard errors in parentheses. Each measure was standardized to have mean 0 and variance 1. Observed individual characteristics include: an indicator for substance abuse, age, age squared, an indicator for incarceration so far, observed accumulated periods of education, legal experience and criminal experience. Periods of sector-specific experience and education are measured in semesters.
*** p<0.01, ** p<0.05, * p<0.1

For crime, there is a positive association with mental health through hostility. According to my estimates, an increase of one standard deviation in hostility is associated with an increase

[14] High school graduation is restricted between ages 17 and 20.

of 2.6 percentage points in the probability of committing an income-related crime and 2.7 percentage points in the probability of committing any type of crime.[15] Therefore, the role of personal capabilities on criminal engagement seems to be independent of the type of crime considered.

What is more interesting perhaps is the role of self-control. Previous studies point to the relation between the lack of self-control and criminal engagement (Gottfredson and Hirschi, 1990; Piquero et al., 2005; Duckworth and Seligman, 2017; Mancino et al., 2016). Evidence of such relation exists when I estimate probit models for criminal engagement (Table B.1). My results show an increase of one standard deviation in self-control is associated with a decrease of 6.8 percentage points for income crime and 6.5 percentage points for any type of crime. My estimates suggest a bigger effect than the one find by Mancino et al. (2016) using PTD as well. However, their definition of individual heterogeneity extends beyond my definition of personal capabilities. In consequence, the bigger effects I find could be a consequence of self-control having to explain more of the variation in crime decisions.

The significant association between self-control and crime decisions does not remain once I estimate the linear models with fixed effects. Coefficients for self-control decrease towards zero when fixed effects are introduced. This suggests that, in the probit specification, self-control is absorbing part of the effect from unobserved fixed characteristics. In the dynamic model from Chapter 4, I also account for the influence of fixed characteristics on choices and find that self-control is an important predictor of criminal engagement in the long-run through its influence over illegal earnings. In addition, in Chapter 4 I account for additional sources of dynamics to study the life-cycle choices and I am able disentangle the long-run implications of changes in self-control on employment and schooling.

The last two columns in Table 3.1 present the estimates for home production and high school graduation. Here, there is a negative association with mental health that works only through depression. In the case of home production, an increase of one standard deviation decreases the probability of being unemployed by 2.7 percentage points. Also, the same increase in depression is associated with a decrease of 1.2 percentage points in the probability of high school graduation. This finding confirms results from previous studies (Fletcher, 2008; Reynolds et al., 2007).[16] My estimates also show a positive and significant association between self-control and high school graduation.[17] This noncognitive skill is influencing the human capital accumulation process.[18]

There are two important conclusions from the previous results. First, different dimensions of mental health problems are influencing different choices. For criminal engagement, the positive association is coming through hostility while for human capital accumulation, the negative association is coming through depression. Second, self-control has a negative association with

[15]When I consider the GSI as a global measure of mental health, such positive and significant association remains only for income crime. According to my estimates (not shown in Table 3.1), an increase of one standard deviation in GSI is associated with an increased of 1.8 percentage points in the probability of committing an income crime and the coefficient is significant at the 10% level.

[16]The magnitude of the association between depression and high school graduation is similar to the one in Fletcher (2008) for his full sample.

[17]Mancino et al. (2016) do not find a significant effect of WAI scores over human capital accumulation decisions. However, they focus on a different margin since their estimates are for enrollment.

[18]These results are in line with Heckman et al. (2006)

crime and positive with high school graduation, but only the latter remains significant after controlling for fixed unobservable characteristics.

## 3.4 Personal Capabilities, Human Capital, and Earnings

In this section, I present evidence about how personal capabilities, sector-specific experience, incarceration records, and education are related to wages and illegal earnings. I use the log potential monthly wages and illegal earnings for the analysis and include only the observations from people who decide to commit a crime or to work as the predominant occupation throughout the semester. Therefore, results are conditional on the participation decision and the classification previously described. The model I estimate is the same as in equation 3.1 and results are presented in Table 3.2. In general, personal capabilities are not significantly associated with wages and illegal earnings. However, coefficients for hostility and self-control are close to the 10% significance level for illegal earnings. In both cases, an increase of one standard deviation is associated with a decrease of 18% in illegal earnings. This evidence suggests that more hostile people may earn less.

Furthermore, people with higher self-control may have lower illegal earnings. My results from Chapter 4 confirm such findings. One compatible explanation is that people with lower self-control are willing to take more risky opportunities earning more money. Another possibility is that lower self-control is associated with more criminal activity on the intensive margin and therefore with higher earnings.

Regarding the returns to experience, incarceration records, and education in the illegal sector, there are three findings. First, there is no evidence of positive and significant returns to education. Even though the coefficient for periods of education is positive, it is too noisy to derive further conclusions.

Second, incarceration records are associated with an increase of more than 100% in illegal earnings. This result is compatible with building criminal capital behind bars (Nguyen et al., 2017; Bayer et al., 2009). Third, there is evidence of returns to criminal experience. This result is in line with findings from Loughran et al. (2013). According to my estimates, the first semester of criminal experience is associated with an increase of 27% in monthly illegal earnings. My estimate for experience squared is negative but insignificant at conventional levels. Together, the results are compatible with the existence of criminal capital. In this version of human capital, skills are created mainly through experience and incarceration. Therefore, previous choices are important predictors for the continuation of criminal behaviour. It is important to mention that I run the same mincer specification on illegal earning while in detention (not shown in the dissertation) and I did not find any evidence on the returns to criminal capital.

The estimates for the returns to experience and education on wages are in the first column of Table 3.2. My estimates show that an additional period of experience increases wages by 6.0%. The returns to education are also positive and significant. An additional period of school increases wages by 5.6%. There is no evidence about the negative and significant effect of incarceration records. This finding is compatible with the fact that the majority of jobs from PTD respondents are low skilled and under-the-table jobs. In these jobs, checks for criminal records are usually not requested by employers (Pager, 2003). To the extent that education does not contribute to building criminal capital and even prevents its accumulation, schooling may

Table 3.2: Mincer regressions with fixed effects for wages and illegal earnings.

| Variable | Wages | Illegal Earnings |
|---|---|---|
| Mental health (previous period) | | |
| Depression | 0.005 | 0.102 |
| | (0.024) | (0.111) |
| Anxiety | -0.019 | -0.008 |
| | -0.023 | (0.131) |
| Somatization | 0.020 | 0.005 |
| | (0.020) | (0.089) |
| Hostility | 0.026 | -0.186 |
| | (0.021) | (0.113) |
| Self-control (previous period) | 0.010 | -0.188 |
| | (0.022) | (0.150) |
| Sector specific experience | 0.060*** | 0.272** |
| | (0.014) | (0.107) |
| Sector specific experience squared | -0.001 | -0.020 |
| | (0.002) | (0.014) |
| Education | 0.056** | 0.093 |
| | (0.025) | (0.185) |
| Incarceration so far | 0.038 | 1.029** |
| | (0.103) | (0.461) |
| Observations | 1,246 | 262 |
| R Square | 0.246 | 0.286 |

Notes: Standard errors in parentheses. Each measure was standardized to have mean 0 and variance 1. Experience is measured in semesters and it ignores pre-survey experience. Dependent variables are log potential monthly wages and illegal earnings. The models are estimated using only the observations from people who decide to commit a crime or to work as their predominant occupation throughout the semester.

*** p<0.01, ** p<0.05, * p<0.1

be an effective deterrent mechanism (Lochner, 2004; Lochner and Moretti, 2004). A detailed discussion about this issue is delayed until Chapter 4.

## 3.5 Changes in Personal Capabilities

In this section, I present evidence on how personal capabilities are related to: (i) the experience of incarceration, (ii) high school graduation, (iii) legal experience, and (iv) criminal experience. I present the analysis into two parts. First, I estimate the following model with fixed effects:

$$h_{it} = \boldsymbol{x}_{it2}\boldsymbol{\beta} + \boldsymbol{d}_{it-1}\boldsymbol{\phi} + s_i + \epsilon_{it}, \tag{3.2}$$

where $h_{it}$ are measures of personal capabilities, $\boldsymbol{x}_{it2}$ are observed individual characteristics such as age, age squared, an indicator for substance abuse problems, accumulated time in detention, accumulated legal and criminal experience, observed periods of school, and an indicator for high school graduation. $\boldsymbol{d}_{it-1}$ is detention status in the previous period,[19] $s_i$ is unobserved heterogeneity, and $\epsilon_{it}$ are idiosyncratic errors. To estimate equation 3.2, I use all periods of data.

Second, I investigate how changes in criminal experience, legal experience, and high school graduation are related to changes in personal capabilities over a longer time frame. For this purpose, I estimate the following model:

$$h_{it} - h_{i0} = (\boldsymbol{x}_{it} - \boldsymbol{x}_{i0})\boldsymbol{\beta} + u_i, \tag{3.3}$$

where $h_{it}-h_{i0}$ is the change in the standardized measures of personal capabilities in the medium-run and in the long-run. In this model, $h_{i0}$ are standardized measures of personal capabilities at the baseline interview and $h_{it}$ are standardized measures two years after the baseline interview for the medium-run and four years after for the long-run. $\boldsymbol{x}_{it} - \boldsymbol{x}_{i0}$ are changes in criminal experience, legal experience, and high school graduation between 0 and t. Finally, $u_i$ is the idiosyncratic error. To estimate equation 3.3, I use two periods of data: the baseline and follow-up interview four for the medium-run and the follow-up interview seven for the long-run.

The model from equation 3.3, allows me to investigate if changes in the independent variables have any effect over changes in personal capabilities over a two-year and a four-year time frame, while the model from equation 3.2 aims to describe the effect of those changes within a period of six months. In addition, I use model 3.3 to focus on the initial periods of observations. During these periods, participants are in the transition from adolescence into adulthood when personality is more malleable (Todd and Zhang, 2019; Almlund et al., 2011) and the onset of mental health problems is more likely to occur (Knapp et al., 2016).

Table 3.3 contains the estimates for the model in equation 3.2. My estimates show the experience of incarceration is associated with more depression, hostility, and GSI. In particular, compared with people who were not incarcerated, those who served time in a juvenile facility are more hostile and depressed and have worse overall mental health. The coefficients associated with serving time in an adult facility are also positive, but for GSI and hostility, they are close to zero and non-significant at conventional levels.

All of these estimates are smaller than the ones from Chapter 2 but the signs of the coefficients are the same. One possible explanation is that, in this chapter, I do not restrict the attention to people without prior incarceration experience. Unlike Chapter 2, I do not restrict the observations to a particular age group. Therefore, results from Table 3.3 can be interpreted as average effects for the PTD sample given the imposition of the linear functional form from equation 3.2. In Chapter 4, I use results similar to the ones in Table 3.3 to describe the law of motion of mental health.[20]

In regards to the accumulation of sector-specific experience, my estimates show that one more year of legal work experience is associated with better mental health. In fact, all the coefficients are negative and they are significant for GSI, hostility, and somatization. For example,

[19]The detention status from an individual $i$ at time $t$ is fully characterized by the following states: (i) no incarceration, (ii) serving time in a juvenile facility, and (iii) serving time in an adult facility. For the regression, indicators of serving time in juvenile and adult facilities are included. Therefore, results should be interpreted to the baseline category "no incarceration".

[20]Results from Table C.5 are similar to the ones from Table 3.3. However, Table C.5 shows an stronger effect of incarceration over mental health and self-control.

Table 3.3: Estimated parameters from regressions with fixed effects on personal capabilities.

| Variables | GSI | Depression | Hostility | Anxiety | Somatization | Self-control |
|---|---|---|---|---|---|---|
| Detention (previous period) | | | | | | |
| Adult Facility | 0.009 | 0.098** | 0.015 | -0.096** | -0.093* | -0.034 |
| | (0.045) | (0.046) | (0.046) | (0.048) | (0.048) | (0.031) |
| Juvenile Facility | 0.089* | 0.097** | 0.095** | 0.063 | -0.010 | -0.020 |
| | (0.047) | (0.048) | (0.048) | (0.050) | (0.049) | (0.033) |
| High school diploma | -0.042 | -0.098 | -0.055 | -0.070 | 0.009 | 0.156*** |
| | (0.061) | (0.064) | (0.064) | (0.066) | (0.014) | (0.043) |
| Legal Experience | -0.029** | -0.012 | -0.032*** | -0.005 | -0.017** | 0.031*** |
| | (0.017) | (0.008) | (0.008) | (0.008) | (0.008) | (0.005) |
| Criminal Experience | 0.005 | 0.041** | -0.023 | 0.020 | -0.013 | -0.000 |
| | (0.019) | (0.017) | (0.017) | (0.018) | (0.017) | (0.012) |
| Observations | 4,033 | 4,033 | 4,033 | 4,033 | 4,033 | 5,615 |
| R-squared | 0.026 | 0.021 | 0.031 | 0.012 | 0.014 | 0.038 |

Notes: Standard errors in parentheses. Observed individual characteristics include age, age squared, an indicator for substance abuse problems, accumulated time in detention, accumulated legal and criminal experience, observed periods of school, and an indicator for high school graduation. Detention status in the previous period is fully characterized by the following states: (i) no incarceration, (ii) serving time in a juvenile facility, and (iii) serving time in an adult facility. For the regression, indicators of serving time in juvenile and adult facilities are included. Therefore, results should be interpreted to the baseline category "no incarceration". All periods of data are used for estimation.
*** p<0.01, ** p<0.05, * p<0.1

one more period of legal work experience is associated with a decrease of 0.029 standard deviations in GSI and of 0.032 standard deviations in hostility. An extra period of employment is associated with an increase in self-control of 0.031 standard deviations. In general, working in the legal sector is associated with better mental health and more self-control.

High school graduation also increases self-control and improves almost all dimensions of mental health. However, my estimates from the latter are not significant at conventional levels. The positive association between high school graduation, employment and self-control is an important result in its own right. Criminologists have argued for years that lack of self-control is a stable personality trait rooted early in life (Nagin and Paternoster, 2000). My results from Table 3.3 show that self-control is malleable and changes during the transition from adolescence into adulthood. I also find evidence that one more period of experience in the criminal sector is associated with more depression.

Tables 3.4 and 3.5 contain the estimates for the model in equation 3.3. During the period that spans the transition from adolescence into adulthood, my estimates show that the medium-run evolution of all dimensions of mental health problems is positively associated with criminal experience. The estimated effects are now in terms of differences of standard deviations. According to my estimates, during a two-year window, a male with one more period of criminal experience becomes more depressed (0.53 std. dev.), anxious (0.48 std. dev.), hostile (0.27 std. dev.), and overall more mentally ill (GSI increases by 0.51 std. dev.). For the long-run, I do not find any significant relationship between changes in personal capabilities.

The previous results are not surprising since criminal engagement during adolescence usually happens in groups where hostile behaviours are reinforced by the influence of peers (Wilson and Herrnstein, 1985). Also, criminal participation is associated with drug and alcohol abuse which is linked with mental health problems (Saffer and Dave, 2002; Fridell et al., 2008;

Schubert et al., 2011; Rajkumar and French, 1997). While changes in personal capabilities associated with employment and high school graduation manifest within a short period, changes associated with criminal participation take time to emerge. In consequence, to understand the implications of the interplay between personal capabilities and choices it is necessary to analyze changes over different time frames.

Table 3.4: Estimated parameters from regressions on the evolution of personal capabilities.

| | GSI | | Depression | | Hostility | |
|---|---|---|---|---|---|---|
| Variables | Medium | Long | Medium | Long | Medium | Long |
| High school diploma | -0.015 | 0.246 | -0.005 | 0.134 | -0.022 | -0.001 |
| | (0.189) | (0.187) | (0.208) | (0.203) | (0.198) | (0.198) |
| Experience | | | | | | |
| Legal | 0.059 | -0.049 | -0.036 | 0.046 | 0.100 | -0.054 |
| | (0.094) | (0.045) | (0.103) | (0.049) | (0.098) | (0.047) |
| Criminal | 0.515*** | 0.034 | 0.537*** | 0.064 | 0.279** | -0.064 |
| | (0.124) | (0.082) | (0.136) | (0.089) | (0.130) | (0.086) |
| Observations | 398 | 272 | 398 | 272 | 398 | 272 |
| R-squared | 0.042 | 0.011 | 0.039 | 0.007 | 0.013 | 0.006 |

Notes: Standard errors in parentheses. The outcome of interest is the difference between the standardized value reported in the baseline interview and the standardized value reported after (i) two years for the medium-run and (ii) four years for the long-run. Two periods of data are used for estimation: the baseline and follow-up interview four for the medium-run and the follow-up interview seven for the long-run.

*** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Estimated parameters from regressions on the evolution of personal capabilities.

| | Anxiety | | Somatization | | Self-control | |
|---|---|---|---|---|---|---|
| Variables | Medium | Long | Medium | Long | Medium | Long |
| High school diploma | 0.052 | 0.216 | -0.018 | -0.047 | 0.138 | 0.152 |
| | (0.206) | (0.202) | (0.219) | (0.204) | (0.116) | (0.121) |
| Experience | | | | | | |
| Legal | 0.001 | -0.067 | 0.014 | -0.062 | -0.003 | 0.021 |
| | (0.102) | (0.048) | (0.109) | (0.049) | (0.059) | (0.029) |
| Criminal | 0.489*** | 0.046 | 0.481*** | 0.015 | 0.022 | -0.038 |
| | (0.135) | (0.088) | (0.144) | (0.089) | (0.083) | (0.052) |
| Observations | 398 | 272 | 398 | 272 | 548 | 431 |
| R-squared | 0.032 | 0.012 | 0.028 | 0.007 | 0.003 | 0.008 |

Notes: Standard errors in parentheses. The outcome of interest is the difference between the standardized value reported in the baseline interview and the standardized value reported after (i) two years for the medium-run and (ii) four years for the long-run. Two periods of data are used for estimation: the baseline and follow-up interview four for the medium-run and the follow-up interview seven for the long-run.

*** p<0.01, ** p<0.05, * p<0.1

These results show that the interplay between choices and personal capabilities is particularly strong for crime and schooling choices. High school graduation both influences and is influenced by self-control. In this case, having more noncognitive skills contributes to the formation of human capital but also higher levels of human capital help to develop noncognitive skills (Todd and Zhang, 2019). The existence of mental health problems increases the likelihood of participating in crime but also criminal engagement (and incarceration) has a negative influence over mental health in the medium-run.

## 3.6 Conclusion

The results from this chapter illustrate the dynamic interaction between choices and personal capabilities. To analyze this dynamic interaction, I use a unique dataset of juvenile offenders: the PTD survey. It is important to stress that my conclusions are based on data coming from adolescents who were found guilty of a serious criminal offence. Although these results shed light on the role of personal capabilities over criminal engagement and human capital formation, they are not necessarily generalizable to the youth population at large. However, I consider these findings important in their own right since people who share characteristics with the PTD participants disproportionately contribute to juvenile crime and incarceration.

My findings suggest that human capital accumulation and criminal engagement are strongly related to personal capabilities, but different capabilities affect different choices. I find evidence that mental health problems influence criminal participation confirming findings from previous studies. Regarding self-control, my estimates show it is an important predictor for human capital accumulation. However, contrary to what previous studies suggest, I do not find a significant association between self-control and crime in the short-run. I come back to this point on Chapter 4 and find that increases in self-control indeed reduces criminal engagement in the long-run mainly through changes in illegal earnings.

In addition, I analyze if previous choices create state dependence through changes in wages and illegal earnings. My findings are compatible with the existence of returns to criminal experience and incarceration records in the illegal sector. I find no evidence of positive and significant returns to education. These findings are compatible with the existence of a particular form of human capital in the criminal sector that is created mainly through experience and incarceration.

I find strong evidence that personal capabilities change over time and are strongly correlated with choices and the consequences of those choices, particularly incarceration. Given that: (i) self-control is a key predictor of human capital accumulation, and (ii) mental health problems interplay with criminal participation and high school graduation, policies designed to improve mental health and to encourage school participation, could become effective crime-fighting strategies over time. I present evidence about this issue in Chapter 4.

# Chapter 4

# Education, Employment and Criminal Capital: Evidence from Juvenile Offenders

## 4.1 Introduction

Over the last several decades, incarceration rates have risen dramatically in many OECD member states. The country that provides the most dramatic example of such an increase is the U.S., where the incarceration rate more than tripled between 1980 and 2012. It has been argued that this increase is, at least in part, the result of using incarceration as the primary crime-fighting strategy (Donohue and Siegelman, 1998). As the population of inmates has continued to rise, policymakers and academics have raised questions about whether incarceration is the most cost-effective intervention to prevent crime.

Arguments against the predominant use of incarceration and the rise of spending on the penitentiary system have prompted a shift in interest to other forms of intervention that focus on changing the opportunity cost of crime.[1] Given the documented negative effect of schooling over criminal engagement and the negative response of crime to higher wages and better job-market opportunities, "alternative" crime-fighting strategies have focused on increasing the incentives to attend school and work, and to develop skills that are rewarded by the labour market.[2]

To design effective policies that discourage criminal participation over the short- and long-

[1]According to Aizer and Doyle (2015), by 2012, federal, state and local expenditures on corrections were expected to exceed \$82 billion per year. Fella and Gallipoli (2014) comment that the average annual cost per prison inmate was \$28,900 in 2008. In addition, some studies have emphasized the negative effects of incarceration on school attainment and employment, and its positive link with re-offending. See for example: Tella and Schargrodsky (2013), Mueller-Smith (2015), and Aizer and Doyle (2015).

[2]Regarding the documented negative relation between schooling and crime, see for example: Lochner (2004), Jacob and Lefgren (2003), and Hjalmarsson (2008). Evidence about the negative relation between crime, wages, and labour-market opportunities is provided in Grogger (1998), Lochner (2004), Imai and Krishna (2004), Mancino (2018), and Raphael et al. (2006).

The benefits of such policies are not only that they reduce crime, but also they increase economic self-sufficiency. These desirable "side-effects" have provided additional reasons to support the implementation of new crime-fighting strategies in the United States. See for example Wald (2016).

run, it is necessary to gain insight into whether changing the opportunity cost of crime can generate different behavioural responses over time. This chapter studies how the dynamic interaction between human capital and criminal capital influences the life-cycle choices of juvenile offenders.[3] Understanding the dynamic interaction between these two types of capital sheds light on why short- and long-run behavioural responses may differ. As my results suggest, when criminal capital accumulates at a faster rate than human capital, temporary pauses in criminal activity may not matter over the long-run since criminal capital can quickly overtake human capital.

There may be other scenarios where the effects of accumulating human capital could be amplified over time. An example of such a scenario is when early investments in education not only stop criminal capital accumulation but also change the opportunity cost of crime in the future because of higher wages. This chapter provides a framework for analyzing when this dynamic interaction either can reinforce the policy incentives to accumulate more human capital or can erode and erase them over time.

To study the influence of human and criminal capital on choices, I develop and estimate a life-cycle model of employment, schooling, and crime from ages 13 to 32. This model allows for the endogenous evolution of both types of capital based on choices made by individuals. I incorporate multiple personal capabilities (mental health, noncognitive skills, and cognitive skills) that influence the evolution of human and criminal capital, and affect the rewards available from schooling and home production. Furthermore, I allow for previous choices to affect the evolution of mental health and noncognitive skills. By including this dynamic dimension of heterogeneity, I control for the fact that some of the personal determinants of criminal participation may change over time.

Finally, I model the effects of incarceration on choices. Individuals can be placed in detention as a consequence of their actions, and such experiences influence the accumulation of human and criminal capital. First, both types of capital are allowed to grow in detention based on whether the individual decides to participate in crime or training programs. Second, the existence of criminal records can directly affect wages and illegal earnings. Third, the evolution of mental health and noncognitive skills is shaped by the experience of punishment.

The estimation of the model is based on individual-level panel data from PTD. As mentioned before, PTD is a multi-site longitudinal dataset that follows a group of juvenile offenders who were convicted of a serious offence in Maricopa County (Phoenix) and Philadelphia County (Philadelphia). This survey contains detailed information about incarceration periods, personal characteristics, school attendance, employment, wages, criminal engagement, and illegal earnings for seven years. The survey also has measures of cognitive skills taken at the time of the baseline interview, as well as repeatedly collected measures of mental health and noncognitive skills.

Three features make my model different from previous work. First, I introduce criminal capital and allow it to evolve endogenously since criminal engagement and incarceration records may increase illegal earnings. This element is absent in most of the previous dynamic models of crime.[4] From an occupational-choice perspective, it is expected that persistent of-

[3] Criminal capital refers to the intangible stock of criminal skills and knowledge that facilitates productive activity in the criminal sector (e.g., to generate illegal earnings).

[4] The existence of returns to criminal experience has been suggested by Loughran et al. (2013), Uggen and Thompson (2003), McCarthy and Hagan (2001), and Mocan et al. (2005).

fenders choose crime because they have a comparative advantage in that sector (Roy, 1951). Criminal capital allows my model to capture the evolution of such a comparative advantage in the criminal sector. Persistent offending is the result of a dynamic selection process in which previous criminal behaviour not only affects human capital and job opportunities but also has a positive influence on illegal earnings. Adding this source of dynamics has important policy implications. As discussed by Lochner (2004), if returns to crime rise with criminal experience, school subsidies could have a sizeable impact because (i) early investments in education encourage further investments in the legal sector; (ii) early investments in education limit the accumulation of criminal experience, and (iii) both effects reinforce one another in discouraging future crime. Not accounting for such dynamic responses may obscure the overall benefits that flow from the policy under consideration.[5]

Second, I allow for choices to depend on personal capabilities and for those capabilities to evolve given individuals' previous choices. In this way, I adjust for the fact that some people become more likely to engage in crime and to go to detention. This dynamic source of heterogeneity has not been included in previous dynamic models of crime. Extensive research in criminology has linked personal attributes to criminal activity. In particular, a strong association has been found between noncognitive skills, especially self-control, and crime (Nagin and Paternoster, 1993; Duckworth and Seligman, 2017).[6]

Some studies also show that self-control is a strong predictor of employment and wealth during adulthood (Moffitt et al., 2011; Heckman et al., 2006). Poor mental health has also been associated with crime (Frank and McGuire, 2010), incarceration (Fagan and Kupchik, 2011), low educational attainment as discussed in Chapter 3,[7] and unemployment (Ettner et al., 1997).

Third, I account for the state dependence generated by incarceration. Some studies have pointed out that detention helps to build criminal capital (Bayer et al., 2009) and has a negative impact on wages due to the existence of a criminal record (Imai and Krishna, 2004; Pager, 2003; Raphael, 2008). In addition, results from Chapter 2 show the negative influence of incarceration on mental health.[8] To the extent that such characteristics are strong predictors of educational attainment and employment, the experience of punishment may induce individuals to accumulate less human capital. As a result, my model accounts for the fact that former inmates may be more likely to recidivate.

The inclusion of these new elements allows me to capture additional sources of dynamics to explain criminal trajectories. In my model, the dynamic selection process of offenders is described not only by the adverse effects of crime on wages and labour market opportunities but also by its positive impact on illegal earnings. I find that my model explains the dynamic pat-

[5]In addition, not accounting for the growth of illegal earnings as a consequence of previous criminal participation may introduce a bias into the estimates of the deterrent effect of detention, schooling and wages.

[6]The role of self-control as a predictor of criminal activity has been discussed extensively among criminologists. As stated by Piquero et al. (2005): "[...] the notion that the difference between offenders and non-offenders lies in their awareness of the concern for the long-term costs of their acts, or self-control, has been supported in extant research, although self-control is not always the strongest or only significant predictor of criminal activity". This chapter provides evidence that self-control is a significant predictor of criminal engagement, but not the most important. Results from Chapter 3 show that even though self-control is not directly associated with criminal engagement, it has an important role on high school graduation and illegal earnings.

[7]Also see Reynolds et al. (2007).

[8]In addition, incarceration has been found to have a negative effect on (i) mental health (Fagan and Kupchik, 2011; Ng et al., 2011), (ii) cognitive function, and (iii) self-control (Umbach et al., 2018).

terns of criminal participation, employment, and school attendance found in the PTD survey. The model's estimates imply that criminal capital accumulates at a faster rate than human capital during the early periods, and that illegal earnings offer a significant premium over wages in the legal sector. I also find that during periods of incarceration, when the evolution of human capital is restricted, criminal capital continues to grow either through the accumulation of criminal experience or through the criminogenic effect of detention. I find evidence of decreasing returns to criminal experience but not to legal experience. According to the estimates, the returns to the first period of criminal experience are more than five times greater than for the first period of experience in the legal sector, and illegal earnings reach their peak after 7.5 years of continuous criminal engagement. As a consequence, criminal capital is an important factor for explaining criminal engagement, especially during the early periods of my model.

Regarding personal capabilities, I find that good mental health increases both illegal earnings and wages. Also, that good mental health increases the value of school but the effect is small. As a result, the existence of mental health problems is not the main driver of criminal engagement in the long-run. I also find that more of self-control is associated with less criminal engagement and lower illegal earnings. As I discussed in Chapter 3, people with more self-control have lower illegal earnings. My results from the dynamic model indicate that an increase of one standard deviation in self-control during the baseline period decreases illegal earnings by 2.4% which generates a long-run decrease in criminal participation of 3.5 percentage points, and the population in detention decreases by 1.3 percentage points. Even though self-control is an important predictor of criminal participation, the main factor that drives decision-making over the long-run is the interaction between both types of capital.

Finally, I find that schooling makes the most significant contribution to human capital accumulation. The returns associated with an additional period of schooling are three times bigger than those associated with an additional period of training, and two times bigger than those from an additional period of experience in the legal sector. As a consequence, any "alternative" crime-fighting strategy implemented during adolescence should be targeted at increasing the incentives to attend school.

Using the estimates from my model, I quantify the effects of two policy interventions: wage and school subsidies. To highlight the importance of the dynamic interaction between human and criminal capital, I simulate the implementation of each policy at two different ages: 16 and 22 years old. There are two key lessons to be drawn from these experiments. First, behavioural responses during and after the subsidy period can be different. Furthermore, when individuals are forward-looking, an anticipated policy could produce results even before implementation. Second, long-run criminal participation depends on the timing of the policy. That is, the same policy implemented at different ages produces different behavioural responses over the long-run.

I find that "early" school subsidies generate investments in education that increase human capital and prevent the accumulation of criminal capital. Since incarceration makes a positive contribution to criminal capital, the deterrent effect of school attendance is also amplified by preventing people from being incarcerated. Over the long-run, this policy increases employment and decreases criminal engagement.

This conclusion is not valid in the case of "early" wage subsidies. In this case, the policy produces only a temporary short-run decrease in criminal participation. During the subsidy period, since employment is the most attractive alternative, some people stop attending school and participating in crime. Such a lower investment in education generates lower salaries in the future. In this case, the rapid accumulation of criminal capital renders futile the human capital accumulated during the subsidy period.

A "late" wage subsidy, on the other hand, increases the number of economically self-sufficient people and reduces criminal engagement over the long-run. Interestingly, this policy also increases investments in education, even before the start of the subsidy period. Since expected wages are higher due to the subsidy, forward-looking agents have an incentive to remain in school and reduce their criminal engagement to increase their probability of employment during the subsidy period.

The rest of the chapter is organized as follows. Section 4.2 describes the structure of the human and criminal capital model. Section 4.3 highlights some of the fundamental characteristics of the data. Section 4.4 discusses the estimation methodology and the identification of the model parameters. Section 4.5 presents the results and discusses them. Section 4.6 discusses the policy experiments. Finally, Section 4.7 presents the conclusions.

## 4.2 A Model of Endogenous Human and Criminal Capital Accumulation

The following is a model with endogenous human and criminal capital that incorporates cognitive skills, mental health and self-control. Individuals can serve time in detention as a consequence of their actions and such experience influences the accumulation of both human and criminal capital. Additionally, choices depend on mental health and self-control, and previous choices also determine the evolution of both factors.

The structure of the model is as follows: each individual has a finite decision horizon beginning at 13 years old and ending forty periods later at age $A = 32.5$. In my model, each period has a length of six months. At each age $a$, an individual chooses among different mutually exclusive and exhaustive alternatives. These alternatives depend on detention status: free or incarcerated. When $i$ is free, the alternatives are: employment ($m = 1$), crime ($m = 2$), school ($m = 3$), or home production ($m = 4$). When $i$ is incarcerated, only two alternatives emerge: crime ($m = 5$) or training ($m = 6$). Note that, during the detention spell, people can not accumulate experience in the legal sector nor years of school. For convenience, I omit the subscript $i$ in the following equations.

Let $d_m(a) = 1$ if alternative $m$ is chosen ($m = 1, ..., 6$) at age $a$, and zero otherwise. Also, let $M^f$ be the set of relevant alternatives when the individual is free, and $M^d$ when the individual is in detention.[9] In the model, each period corresponds to a semester. The rewards per period, at any age $a$, are given by

$$R(a) = \sum_{m=1}^{6} R_m(a)d_m(a), \tag{4.1}$$

where $R_m(a)$ is the reward per period associated with the $m$th alternative. These rewards contain all the benefits and costs associated with each alternative.

[9] $m \in M^f$ for $m = 1, ..., 4$ and $m \in M^d$ for $m = 5, 6$.

### 4.2.1 Rewards

#### Rewards when Not in Detention

**Employment ($d_1(a) = 1$):** The current period reward for working is the wage, $w_1(a)$, and nonpecuniary benefits ($\gamma_{1,1}$):

$$R_1(a) = w_1(a) + \gamma_{1,1}. \tag{4.2}$$

The wage is the product of the rental price ($r_1$) times the number of human capital units possessed by the individual, $hc_1(a)$. Then

$$w_1(a) = r_1 \cdot hc_1(a). \tag{4.3}$$

To identify the constants in the reward functions, I normalize $\gamma_{1,1} = 0$ and $r_1 = 1$. Therefore, all of the other constants should be interpreted as relative to $\gamma_{1,1}$. More details about identification are given in Section 4.4.2.

In a standard human capital formulation, the level of skills accumulated up to any age depends on the number of periods of schooling ($s(a)$), and work experience ($x_1(a)$), which typically takes a quadratic form (Mincer, 1958). In this model, I also include an observed measure of fixed cognitive ability ($c$), self-control ($nc(a)$), and mental health ($mh(a)$). Mental health has been related to productivity in the workplace (Ettner et al., 1997; Lerner and Henke, 2008) and, as discussed in Chapter 3, self-control is a key predictor of educational outcomes.[10]

Incarceration enters $hc_1(a)$ in different ways. First, it affects the evolution of $nc(a)$ and $mh(a)$ as discussed in Subsection 4.2.2. Second, in detention people can acquire skills ($trg(a)$) by participating in training programs. Third, criminal records may impact wages due to stigma (Pager, 2003; Raphael et al., 2006). To account for this channel, I include $q(a)$ which takes the value of 1 when an individual has been in detention at least once up to age $a$.

Finally, I allow skills to depreciate in order to capture the potential reduction in wages for people who were not working in the previous period ($d_1(a-1) = 0$). Letting $hc_1$ be the number of human capital units possessed at age $a$, $e_1(13)$ the skill endowment at age 13,[11] and $\epsilon_1(a)$ a skill technology shock. I define $hc_1$ as:

$$\begin{aligned} hc_1(a) = \exp\{&e_1(13) + \alpha_{1,1} \cdot s(a) + \alpha_{1,2} \cdot trg(a) + \alpha_{1,3} \cdot x_1(a) + \alpha_{1,4} \cdot x_1^2(a) + \alpha_{1,5} \cdot q(a) \\ &+ \alpha_{1,6} \cdot c + \alpha_{1,7} \cdot mh(a) + \alpha_{1,8} \cdot nc(a) - \alpha_{1,9} \cdot (1 - d_1(a-1)) + \epsilon_1(a)\}. \end{aligned} \tag{4.4}$$

This specification leads to a (ln) wage equation in which the constant term is $\ln(r_1) + e_1(13)$, the sum of the ln rental price and the age 13 skill endowment.

**Crime ($d_2(a) = 1$):** The contemporaneous rewards for criminal engagement are given by:

$$R_2(a) = w_2(a) + \gamma_{2,1} + \gamma_{2,2} \cdot nc(a), \tag{4.5}$$

where $w_2(a)$ are the illegal earnings:

$$w_2(a) = r_2 \cdot hc_2(a). \tag{4.6}$$

[10] In addition, see Heckman et al. (2006); Moffitt et al. (2011).

[11] I estimate a version of the model without heterogeneity in initial endowments and so $e_1(13)$ is the same for everybody.

In this specification, $\gamma_{2,1} + \gamma_{2,2} \cdot nc(a)$ are the non-monetary benefits of crime that may change with $nc(a)$. The idea is to account for the fact that people with lower self-control may tend to pursue short-term, immediate rewards and gratification, neglecting the long-term consequences of the criminal activity (Piquero et al., 2005). Therefore, lower $nc(a)$ may imply bigger non-monetary benefits.

Illegal earnings are the product of the rental price ($r_2$) times the criminal capital up to age $a$, $hc_2(a)$. The production function of criminal capital depends on the skill endowment at age 13 ($e_2(13)$), criminal experience ($x_2$), skill technology shock ($\epsilon_2(a)$), $nc(a)$, $mh(a)$, and $c$. I also allow skills to depreciate (when $d_2(a-1) = 0$) and to change due to criminal records (when $q(a) = 1$). The specification for $hc_2(a)$ is:

$$\begin{aligned} hc_2(a) = \exp\{&e_2(13) + \alpha_{2,1} \cdot x_2(a) + \alpha_{2,2} \cdot x_2^2(a) + \alpha_{2,3} \cdot q(a) + \alpha_{2,4} \cdot c + \alpha_{2,5} \cdot mh(a) \\ &+ \alpha_{2,6} \cdot nc(a) - \alpha_{2,7} \cdot (1 - d_2(a-1)) + \epsilon_2(a)\}. \end{aligned} \tag{4.7}$$

This specification leads to a (ln) earnings equation in which the constant term is $\ln(r_2) + e_2(13)$, the sum of the ln rental price and the age 13 skill endowment. To identify $e_2(13)$, I also normalize $r_2 = 1$. Finally, the expected cost of crime is given by the incarceration spell which I discuss later. I assume individuals, with some probability, start the incarceration spell in the next period and therefore, the costs of crime are not part of the contemporaneous reward function.

**School ($d_3(a) = 1$):** The reward for school attendance, which is denominated in dollars, is a function of the net value of attending school.

This value is defined by a fixed, non observable, endowment ($e_3(13)$) (i.e. net taste for school), a fixed observable cognitive component ($c$), and several variable components that enter linearly: age ($a$), mental health ($mh(a)$), noncognitive skills ($nc(a)$), previous school ($d_3(a-1)$) and training ($d_6(a-1)$), and a random element ($\epsilon_3(a)$). The contemporaneous reward takes the following form:

$$\begin{aligned} R_3(a) = &e_3(13) + \alpha_{3,1}\mathbb{1}_{(16\leq a\leq 18)} + \alpha_{3,2}\mathbb{1}_{(18<a)} + \alpha_{3,3} \cdot a + \alpha_{3,4}\mathbb{1}_{(16\leq a\leq 18)} \cdot a \\ &+ \alpha_{3,5}\mathbb{1}_{(18<a)} \cdot a + \alpha_{3,6} \cdot c + \alpha_{3,7} \cdot mh(a) + \alpha_{3,8} \cdot nc(a) \\ &- \alpha_{3,9} \cdot (1 - d_3(a-1) - d_6(a-1)) + \epsilon_3(a). \end{aligned} \tag{4.8}$$

The coefficients associated with $c$, $mh(a)$, and $nc(a)$, measure how rewards change with these characteristics. For example, an individual with higher cognitive ability and self-control may find it more enjoyable to attend school. Similarly, lower mental health could make attending school an unpleasant experience, and therefore the net value should be lower. Finally, I include re-entry costs ($\alpha_{3,9}$) to capture potential knowledge depreciation.[12]

It seems reasonable to think that, when attendance is not compulsory (i.e. after age 16) school may have a consumption value that changes with age, especially after the age of high

[12] If school attendance is not continuous, $i$ may have to increase his effort to compensate for having forgotten what he once learned. Also, $i$ could dislike attending school with different classmates from a younger cohort. In both cases, the consumption value is expected to decrease.

school graduation when the majority of people in my sample is working. In my specification, the age dummies and their interactions with age capture such differences.[13]

**Home Production ($d_4(a) = 1$):** The reward for remaining at home, which is also denominated in dollars, depends on a fixed, non observable, endowment $e_4(13)$ (i.e. net taste for leisure), age, mental health, cognitive ability and self-control, and a random component ($\epsilon_4(a)$). I include $nc(a)$ and $c$ as arguments of the reward function to account for the fact that people with lower levels of both measures may find it more rewarding to stay at home and not face the challenges associated with attending school or being employed. Besides, $R_4(a)$ depends on $mh(a)$ since the value of staying at home could be higher for people with low mental health. The reward for home production takes the following functional form:

$$\begin{aligned} R_4(a) =& e_4(13) + \alpha_{4,1}\mathbb{1}_{(16\leq a\leq 18)} + \alpha_{4,2}\mathbb{1}_{(18<a)} + \alpha_{4,3}\cdot a + \alpha_{4,4}\mathbb{1}_{(16\leq a\leq 18)}\cdot a \\ &+ \alpha_{4,5}\mathbb{1}_{(18<a)}\cdot a + \alpha_{4,6}\cdot c + \alpha_{4,7}\cdot mh(a) + \alpha_{4,8}\cdot nc(a) + \epsilon_4(a). \end{aligned} \tag{4.9}$$

For this option, $\alpha_{4,1}$, $\alpha_{4,2}$, and $\alpha_{4,3}$ reflect changes in the payoff that come with age. Those changes can be caused, for example, by welfare transfers which individuals are more likely to obtain when they grow older.[14] In this chapter, I use home production and unemployment interchangeably.

### In Detention Rewards

While in detention, offenders continue reporting criminal earnings and accumulating years of training. As a consequence, I model these choices because they have an impact on the accumulation of human and criminal capital.

**Crime in Detention ($d_5(a) = 1$):** The contemporaneous reward for criminal engagement is given by:

$$R_5(a) = \exp\{e_5(13) + \epsilon_5(a)\} + \gamma_{5,1} + \gamma_{5,2}\cdot nc(a), \tag{4.10}$$

where $\exp\{e_5(13) + \epsilon_5(a)\}$ are the monetary benefits. While in detention, earnings only depend on the initial endowment of criminal ability ($e_5(13)$) and a random element ($\epsilon_5$). I choose this specification for two reasons. First, the nature of earnings is different. For example, individuals may decide to participate in crime as an exchange for protection or goods. Other reasons could be gaining respect and acceptance from other inmates. In all these cases, there are no monetary returns to criminal experience. Second, in the reduced form analysis of illegal earnings in detention (not shown in the dissertation) I did not find evidence on the returns to criminal capital.

[13] The age variables also help the model to rationalize changes in school participation and home production observed after age 16 given the absence of discrete types in the estimation.

[14] The data show a positive association between non-salary monetary transfers and age. These non-salary transfers would affect the available income for people who decide to stay at home and consequently their valuation of this option. For example, $i$ may get a different payoff for staying at home when he is 17 years old and did not receive any welfare transfers, because he may be in the family group of his parents.

$\gamma_{5,1}$ represents the net nonmonetary benefits for this alternative, i.e., the baseline consumption minus the fixed cost of incarceration. Such cost refers to the psychological pain, measured in dollars, for being deprived of freedom. Also, I allow the nonmonetary benefits to change with $nc(a)$ for the same reasons I mentioned for $d_2(a) = 1$. For identification purposes, I set $\gamma_{5,1} = 0$ and so $e_6(13)$ should be interpreted as relative to $\gamma_{5,1}$. More details about identification are given in Section 4.4.2.

**Training ($d_6(a) = 1$):** The reward for training in detention has a similar specification as attending school.[15] In particular:

$$\begin{aligned} R_6(a) = & e_6(13) + \alpha_{6,1}\mathbb{1}_{(16\leq a\leq 18)} + \alpha_{6,2}\mathbb{1}_{(18<a)} + \alpha_{6,3}\cdot a + \alpha_{6,4}\mathbb{1}_{(16\leq a\leq 18)}\cdot a \\ & + \alpha_{6,5}\mathbb{1}_{(18<a)}\cdot a + \alpha_{6,6}\cdot c + \alpha_{6,7}\cdot mh(a) + \alpha_{6,8}\cdot nc(a) \\ & + \alpha_{6,9}\cdot(1 - d_3(a-1) - d_6(a-1)) + \epsilon_6(a). \end{aligned} \tag{4.11}$$

In this specification, $e_6(13)$ represents the net value of training, i.e., the consumption value minus the fixed cost of incarceration. All the other components in the reward function are included for the same reasons as in the reward for school.

### 4.2.2 State Space

#### Law of Motion

The state variables such as schooling and sector-specific experience evolve in a deterministic manner that is (conditionally) independent of the shocks. In particular, the evolution of (i) education is $s(a+1) = s(a)+d_3(a)$, (ii) training is $trg(a+1) = trg(a)+d_6(a)$, (iii) legal experience is $x_1(a+1) = x_1(a) + d_1(a)$, and (iv) criminal experience is $x_2(a+1) = x_2(a) + d_2(a) + d_5(a)$.

Finally, the evolution of $mh(a)$ and $nc(a)$ depends on the detention status from the previous period, $mh(a-1)$ and $nc(a-1)$, respectively, and shocks.

$$\begin{aligned} mh(a) = & \alpha_{mh,0} + \alpha_{mh,1}\cdot mh(a-1) - \alpha_{mh,2}\cdot j(a-1)\cdot(1 - f(a-1)) \\ & - \alpha_{mh,3}\cdot(1 - j(a-1))\cdot(1 - f(a-1)) + \alpha_{mh,4}\cdot f(a-1) + \xi_{mh}(a) \end{aligned} \tag{4.12}$$

$$\begin{aligned} nc(a) = & \alpha_{nc,0} + \alpha_{nc,1}\cdot nc(a-1) - \alpha_{nc,2}\cdot j(a-1)\cdot(1 - f(a-1)) \\ & - \alpha_{nc,3}\cdot(1 - j(a-1))\cdot(1 - f(a-1)) + \alpha_{nc,4}\cdot f(a-1) + \xi_{nc}(a) \end{aligned} \tag{4.13}$$

In equations 4.12 and 4.13, $\alpha_{mh,0}$ and $\alpha_{nc,0}$ capture the average effect of all factors that are not included in the model. The detention status is given by the combination of placement status ($f(a)$) and the facility type ($j(a)$).[16] Therefore, $\alpha_{nc,2}$ and $\alpha_{mh,2}$ measure the depreciation rate associated with serving time in a juvenile facility, and $\alpha_{nc,3}$ and $\alpha_{mh,3}$ the depreciation

[15] It is important to mention that the majority of people (94.4%) below 16 years old choose this alternative. This could be due to institutions with the need to fulfill requirements for compulsory school attendance. In any case, this fact supports the idea that consumption value may change with age.

[16] If $i$ is free at age $a$, then $f(a) = 1$. When $i$ is in detention $f(a) = 0$ and $j(a)$ takes the value of 1 for juvenile facility and 0 for adult facility.

rate associated with adult facilities. Similarly, $\alpha_{nc,4}$ and $\alpha_{mh,4}$ measure the change in both characteristics when $i$ is not in detention. I include a facility type effect since evidence from Chapter 3 suggests that adult and juvenile placements have different consequences on mental health and the evolution of noncognitive skills.[17] To reduce the number of parameters to be estimated within the model, I estimate the parameters from 4.12 and 4.13 outside the model for reasons explained in Section 4.4.3. The estimates from 4.12 and 4.13 are similar to the ones in Section 3.5. For the sake of simplicity, the law of motion does not account for the existence of heterogeneous effects of incarceration similar to the ones discussed in Chapter 2.

#### Initial Endowments and Random Elements

Initial conditions at age 13 are given by the accumulated number of periods of schooling ($s(13) = 16$), training ($trg(13) = 0$), work and criminal experience ($x_1(13) = x_2(13) = 0$), criminal records ($q(13) = 0$), and the baseline level of mental health ($mh(13)$) and self-control ($nc(13)$).[18]

In addition, each $i$ gets an initial endowment vector $\boldsymbol{e(13)}$={$e_1(13)$, $e_2(13)$, $e_3(13)$, $e_4(13)$, $e_5(13)$, $e_6(13)$}.[19] Define the experience vector $\boldsymbol{x(a)} = \{x_1(a), x_2(a)\}$, the schooling vector $\boldsymbol{s(a)} = \{s(a), trg(a)\}$, the personal capabilities vector $\boldsymbol{h(a)}$ ={$mh(a)$, $nc(a)$, $c$}, and the vector with the decisions from the previous period $\boldsymbol{d(a-1)}$.

Also define, the detention status vector $\boldsymbol{\rho(a)} = \{l(a), l(a-1), j(a), j(a-1), q(a)\}$, where $l(a)$ is the number of remaining periods in detention, $j(a)$ indicates if the individual is serving time in a juvenile facility, and $q(a)$ takes the value of 1 after the first incarceration spell at age $a$. Finally, define the reward shock vector $\boldsymbol{\epsilon(a)}$={$\epsilon_1(a)$, $\epsilon_2(a)$, $\epsilon_3(a)$, $\epsilon_4(a)$, $\epsilon_5(a)$, $\epsilon_6(a)$}, and the personal capabilities shock vector $\boldsymbol{\xi(a)} = \{\xi_{mh}(a), \xi_{nc}(a)\}$. Further, I denote $\boldsymbol{\Phi(a)}$={ $\boldsymbol{e(13)}$, $\boldsymbol{x(a)}$, $\boldsymbol{s(a)}$, $\boldsymbol{h(a)}$, $\boldsymbol{d(a-1)}$, $\boldsymbol{\rho(a)}$, $\boldsymbol{\epsilon(a)}$, $\boldsymbol{\xi(a)}$ } to be the state space vector at age $a$.

To close the model, reward shocks are assumed to be jointly normal ($\mathcal{N}\{0, \Omega\}$), and serially uncorrelated. $\boldsymbol{\xi(a)}$ has a nonparametric distribution that is estimated from the data. It is also assumed to be independent of $\boldsymbol{\epsilon(a)}$. Finally, I assume $\boldsymbol{\xi(a)}$ and $\boldsymbol{\epsilon(a)}$ are independent of all past state variables.

### 4.2.3 The Recursive Problem

In the model, time is discrete and finite with $a = 13, ..., A$. At the beginning of time, $i$ is not in detention ($f(13) = 1$). The individual gets an endowment $\boldsymbol{e(13)}$ and $\boldsymbol{h(13)}$. The individual observes all of these elements, but they are not observed by the researcher. Also, $i$ starts with $\boldsymbol{x(13)}$, and $\boldsymbol{s(13)}$ and draws the random shocks $\boldsymbol{\epsilon(13)}$ and $\boldsymbol{\xi(13)}$. Then, $i$ uses the shocks to calculate the contemporaneous rewards and thus the alternative specific value functions for

[17] Also see Fagan and Kupchik (2011). Based on the PTD data, Mulvey et al. (2007) report that the number of services available to convicts is generally lower in jails and prisons, which are the typical examples of adult detention facilities. Such services normally include drug and alcohol treatment, sessions with psychologist or psychiatrist, group therapy, treatment on mental health unit, and anger management or social skills training.

[18] I assume that $x_2(13) = trg(13) = q(13) = 0$ because 89% of my sample reported the age of the first crime to be 13 years old and over. However, my model can not accommodate the fact that people entering the PTD survey at different ages could be systematically different in terms of unobservable characteristics.

[19] The model is estimated without this dimension of heterogeneity. Therefore, I assume that $\boldsymbol{e(13)}$ is the same for everybody.

four different options: (i) employment, (ii) crime, (iii) school, and (iv) home production. Then, $i$ chooses the alternative that yields the highest value, and the state space is updated accordingly. It is important to clarify that $i$ observes all the realized shocks at $a$, but the researcher does not.

If the individual chooses crime, he goes to detention with probability $\pi_2$, after collecting the rewards. If the individual does not go to detention, next period, he chooses between the same alternatives. If the individual goes to detention, the incarceration spell starts in $a + 1$. When someone first enters detention there are two elements randomly drawn. First, $i$ gets a random draw of the facility type where he has to spend time: juvenile facility ($j(a + 1) = 1$) or adult facility ($j(a + 1) = 0$). I assume that $i$ does not change facilities while incarcerated. Second, $i$ gets a random draw of the penalty length $l(a+1)$ which decreases deterministically after that.[20] All probabilities are assumed to be known by $i$. Details about the distributions of $j(a + 1)$ and $l(a + 1)$ are presented in Section 4.4.3.

Since my model only focuses on criminal activities that generate income, some criminal engagement may happen outside my model. Also, incarceration may happen with some delay after a crime takes place (e.g. the time of the trial). To accommodate these situations, I assume an individual may go to detention when $m = 1, 3, 4$ is chosen. Therefore, I also include probabilities $\pi_m$, for $m \neq 2$, which are time-invariant and estimated outside the model.[21]

Once in detention, the individual draws the random shocks from $\boldsymbol{\epsilon(a)}$ and $\boldsymbol{\xi(a)}$. Then, the shocks are used to calculate the alternative specific value functions for two different options: (i) crime and (ii) training. The alternative that yields the highest value is chosen. This process is repeated in each period during the incarceration spell.

For the last decision period ($A$), I assume the individual is not in detention and chooses the alternative that yields the highest payoff.[22]

At any age, the individual's objective is to maximize the expected present value of remaining lifetime rewards. Defining the value function $V(\boldsymbol{\Phi(a)}, a)$ to be the maximal expected present value of lifetime rewards at age $a$ given $\boldsymbol{\Phi(a)}$ and discount factor $\beta$, the Bellman equation when $i$ is free is given by:

$$V^f(\boldsymbol{\Phi(a)}, a) = \max_{m \in M^f}\{V^f_m(\boldsymbol{\Phi(a)}, a)\}, \tag{4.14}$$

where $V^f_m(\boldsymbol{\Phi(a)}, a)$, the alternative-specific value functions for $a < A$, are given by:

$$\begin{aligned} V^f_m(\boldsymbol{\Phi(a)}, a) =& R_m(\boldsymbol{\Phi(a)}, a) + \beta\mathbb{E}[\pi_m \cdot V^d(\boldsymbol{\Phi(a+1)}, a + 1|\boldsymbol{\Phi(a)})] \\ & + \beta\mathbb{E}[(1 - \pi_m) \cdot V^f(\boldsymbol{\Phi(a+1)}, a + 1)|\boldsymbol{\Phi(a)})]. \end{aligned} \tag{4.15}$$

The Bellman equation associated with being in detention is given by:

$$V^d(\boldsymbol{\Phi(a)}, a) = \max_{m \in M^d}\{V^d_m(\boldsymbol{\Phi(a)}, a)\}, \tag{4.16}$$

and the alternative-specific value functions are defined as follows:

[20] The law of motion for $l(a + 1)$ is given by $l(a + 1) = l(a) - 1$ for $l(a) \geq 0$. Therefore, there is no random release from incarceration.

[21] These are the incarceration frequencies, conditional on last period choices, that I observe in the data.

[22] The maximum penalty $i$ can get is $l_{max}(a) = A - (a + 1)$. Therefore $l_{max}(a) \to 0$ as $a \to A$.

$$V_m^d(\boldsymbol{\Phi(a)}, a) = \begin{cases} R_m(\boldsymbol{\Phi(a)}, a) + \beta\mathbb{E}[V^d(\boldsymbol{\Phi(a+1)}, a+1|\boldsymbol{\Phi(a)}) & \text{if } l(a+1) > 0 \\ R_m(\boldsymbol{\Phi(a)}, a) + \beta\mathbb{E}[V^f(\boldsymbol{\Phi(a+1)}, a+1|\boldsymbol{\Phi(a)}) & \text{if } l(a+1) = 0. \end{cases} \tag{4.17}$$

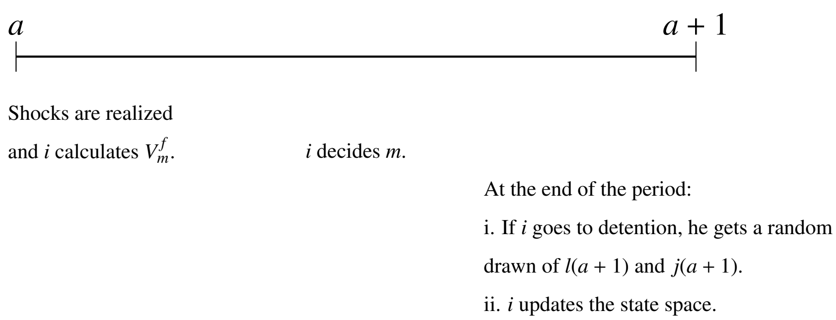
For the last period, the alternative-specific value functions are:

$$V_m^f(\boldsymbol{\Phi(A)}, A) = R_m(\boldsymbol{\Phi(A)}, A), \tag{4.18}$$

where $R_m$ is the alternative-specific reward function previously defined (equations 4.2-4.9).

The expectation is taken over the distribution of the random components of $\boldsymbol{\Phi(a+1)}$. That is, over the unconditional distribution of $\boldsymbol{\epsilon(a+1)}$ and $\boldsymbol{\xi(a+1)}$. When $i$ is free, the expectation is also taken over the conditional distributions of $j(a+1)$ and $l(a+1)$.[23] Figure 4.1 represents the recursive decision process previously described when $i$ is not in detention.

Figure 4.1: Timing in the Model (Not in Detention)




## 4.3 Data

The data were described in detail in Sections 2.2 and 3.2. As indicated before, the sample I use for estimation consists of 779 individuals who were less than 18 years old at the time of the baseline interview. For all of these individuals, I have completed 14 biannual histories giving 10,906 observations in the data set. In this section, I present some descriptive statistics from that subsample. I argue that some features of the data are compatible with a human capital framework to study the choices of juvenile offenders.

Table 4.1 reports demographic and descriptive statistics from my sample. At the baseline interview, the average age of respondents was 16.5 years, 45% of them were in detention, and the same proportion had prior experience of being placed in a residential facility. Even though all participants were under 18 years old, approximately 25% of them were placed in adult facilities. There is a significant representation of minorities in the sample with 36% of the participants being black and 35% hispanic. On average, each participant had 1.31 detention spells during the time covered by the survey, and their average duration was 1.05 semesters.

[23]The probabilities of $j(a+1)$ and $l(a+1)$ depend on some observable characteristics as I explain in Section 4.4.3.

Regarding personal capabilities from participants, it is important to highlight two facts. First, the average cognitive ability is 84.5 which is below the population average of 100. Second, the average mental health in my sample, at the baseline interview, is lower than in the "normative group", implying that people in my sample have poorer mental health than subjects without mental health problems.[24] Throughout this chapter, I work with the overall measure of mental health provided by the GSI. This choice is made mainly because, given the complexity of the model, it is convenient to specify only one law of motion for this attribute and not one for each of the dimensions analyzed in the previous chapters.

Some features of the data are compatible with a human capital framework to study life-cycle choices. In particular, I highlight the following. First, school attendance decreases with age. Second, employment increases with age. Third, the age-crime profile does not exhibit a sharp decline after adolescence. Fourth, wages and illegal earnings increase over time. Fifth, choices exhibit some degree of persistence.

Table 4.1: Sample Description

| Variable | Average | Std. Dev. |
|---|---|---|
| Philadelphia | 0.45 | (0.49) |
| Male | 0.83 | (0.36) |
| Age | 16.53 | (1.00) |
| Black | 0.36 | (0.48) |
| White | 0.24 | (0.42) |
| Hispanic | 0.35 | (0.47) |
| Other | 0.05 | (0.19) |
| Age first crime | 14.87 | (1.56) |
| Duration detention spell | 3.27 | (2.92) |
| Number of detention spells | 1.31 | (1.12) |
| Measures at the baseline | | |
| Ever in Detention | 0.45 | (0.49) |
| Juvenile Facility | 0.75 | (0.43) |
| Cognitive Skills | 84.52 | (13.02) |
| Mental Health | 0.54 | (0.44) |
| Noncognitive Skills | 2.82 | (0.80) |
| People | 779 | |

Table 4.2 shows the proportion of people choosing each alternative by age. As expected, the proportion of people in school declines with age, with the sharpest decline around 16 years old, which is close to the age limit for compulsory attendance in Arizona and Pennsylvania.[25] Interestingly, school attendance decreases steadily up to age 20, before the typical age of college graduation, confirming that disadvantaged youth acquire fewer years of education.

An increase in employment accompanies the decrease in school attendance. In the legal sector, participation is around 7% by age 15 and 57% by age 20. After age 20, it increases to over 60% for all ages. In the criminal sector, the pattern is somewhat similar. By age 15 the

[24] According to Derogatis and Melisaratos (1983), one way to interpret the GSI score is to compare it against the mean from a group of subjects without mental health problems ("normative group"). The mean value for the "normative group" in Derogatis and Melisaratos (1983) is 0.30 with a standard deviation of 0.31.

[25] In fact, the compulsory age for school attendance is 16 years old in Arizona and 17 years old in Pennsylvania.

participation rate is around 2%. Then, it increases steadily up to age 20 where it reaches its peak of about 8%. After that, participation in crime decreases slightly to around 6% suggesting an age-crime profile which is compatible with the existence of positive but decreasing returns to criminal experience.[26] This decrease is less pronounced than the one found in previous studies (Quetelet, 1984; Leung, 1994; Hjalmarsson and Bindler, 2017), although my definition of criminal participation is not the same.[27] The evolution of criminal participation in detention is less clear since it decreases from age 15 to age 18, then increases until age 21. For the final years, it fluctuates between 13% and 8%. Such a pattern is not compatible with the existence of returns to criminal experience in detention.

Table 4.2: Choice Distribution (in percentage)

| Age | Choice | | | | | |
|---|---|---|---|---|---|---|
| | Freedom | | | | Detention | |
| | Employment | Crime | School | Home | Crime in detention | Training |
| 16 | 0.092 | 0.046 | 0.803 | 0.059 | 0.040 | 0.960 |
| 16.5 | 0.15 | 0.053 | 0.732 | 0.065 | 0.035 | 0.965 |
| 17 | 0.237 | 0.051 | 0.615 | 0.098 | 0.051 | 0.949 |
| 17.5 | 0.366 | 0.082 | 0.434 | 0.119 | 0.045 | 0.955 |
| 18 | 0.409 | 0.091 | 0.325 | 0.176 | 0.102 | 0.898 |
| 18.5 | 0.507 | 0.077 | 0.217 | 0.199 | 0.097 | 0.903 |
| 19 | 0.531 | 0.076 | 0.147 | 0.245 | 0.106 | 0.894 |
| 19.5 | 0.55 | 0.079 | 0.124 | 0.247 | 0.101 | 0.899 |
| 20 | 0.559 | 0.08 | 0.100 | 0.262 | 0.106 | 0.894 |
| 20.5 | 0.583 | 0.078 | 0.094 | 0.245 | 0.094 | 0.906 |
| 21 | 0.589 | 0.058 | 0.087 | 0.265 | 0.126 | 0.874 |
| 21.5 | 0.632 | 0.064 | 0.079 | 0.225 | 0.143 | 0.857 |
| 22 | 0.603 | 0.056 | 0.085 | 0.255 | 0.088 | 0.912 |
| 22.5 | 0.653 | 0.062 | 0.079 | 0.206 | 0.089 | 0.911 |
| 23 | 0.602 | 0.076 | 0.08 | 0.241 | 0.104 | 0.896 |

Table 4.3 shows the average wages and illegal earnings for ages 15 to 24. In general terms, wages and illegal earnings increase with age, although illegal earnings decrease significantly at ages 23 and 24, which is compatible with the existence of decreasing returns to criminal experience. It is important to highlight the difference in the average and the growth rate between the two sectors. By age 15, the average monthly wage in the criminal sector is almost two times the wage in the legal sector. Between ages 15 and 22 wages almost doubled in the legal sector and tripled in the criminal sector. Such a difference in levels and growth rates, suggests the

[26]This pattern of criminal participation is robust to alternative definitions of crime. For example, If I classify as doing crime for the whole semester someone who engages in crime at least during one month, participation also grows with age up to age 17 and is around 4 percent higher than the one presented here. After that, criminal participation under both definitions is practically the same. Since both definitions produce the same crime-age, it is unlikely that my results are driven by the definition of crime adopted.

[27]To study the relationship between age and crime Quetelet (1984); Leung (1994); Hjalmarsson and Bindler (2017) use the variation in the arrest rate by age.

existence of a premium for criminal activities.

Finally, Table 4.4 shows one-period transition rates between choices when individuals are free in two consecutive periods. It is important to notice the high degree of persistence for employment in the legal sector. Approximately 78% of those who were working in the legal sector chose the same option for the following period. For crime, the situation is similar. Around 60% of those working in the criminal sector continue with that activity the following period.

Regarding school, persistence is also around 60%, and the majority of people leaving school chose to work (22.8%). Such a transition is expected if years of school have value only for employment. Finally, more than 60% of those who chose home production continue with that alternative the following period, suggesting significant state dependence. The observed persistence in choices can be explained by the existence of returns to experience, depreciation of skills, and re-entry costs to school. These are key features of my model.

Table 4.3: Monthly Potential Wages and Illegal Earnings

| Age | Employment | | Crime | |
|---|---|---|---|---|
| 16 | 1409.93 | (53) | 3074.60 | (20) |
| 17 | 1309.11 | (216) | 3569.00 | (50) |
| 18 | 1390.91 | (452) | 4220.24 | (95) |
| 19 | 1603.20 | (610) | 3910.65 | (81) |
| 20 | 1699.51 | (628) | 5664.94 | (83) |
| 21 | 2005.94 | (645) | 6095.69 | (62) |
| 22 | 2062.55 | (497) | 6391.97 | (49) |
| 23 | 2167.39 | (303) | 3703.03 | (33) |
| 24 | 2619.46 | (80) | 3850.00 | (8) |

Notes:
Number of observations in parentheses

Table 4.4: Transition Matrix: Choices in Freedom

| | Choice (t) | | | |
|---|---|---|---|---|
| Choice (t-1) | Employment | Crime | School | Home |
| Employment | 0.788 | 0.028 | 0.069 | 0.116 |
| Crime | 0.183 | 0.599 | 0.083 | 0.135 |
| School | 0.228 | 0.039 | 0.624 | 0.109 |
| Home | 0.284 | 0.037 | 0.081 | 0.597 |

## 4.4 Estimation and Identification

### 4.4.1 Estimation

This model is estimated by indirect inference (Gourieroux et al., 1993). Let the set of structural parameters be $\Theta$. The method involves simulating data from the model (given a hypothesized value of $\Theta$) and choosing the estimator $\tilde{\Theta}$ to make the simulated data match the actual data as closely as possible according to some criterion that involves a set of auxiliary models. The idea is to minimize the distance between the auxiliary models estimated from the data $\beta$ and the same models estimated from the simulated data $\hat{\beta}(\tilde{\Theta})$ according to some metric. The following objective function represents such a metric:

$$\Phi(\tilde{\Theta}) = (\beta - \hat{\beta}(\tilde{\Theta}))W^{-1}(\beta - \hat{\beta}(\tilde{\Theta})),$$

where $W^{-1}$ is the weighting matrix. I use a diagonal weighting matrix during estimation, where each diagonal element of $W^{-1}$ is the inverse of the variance of the corresponding moment.

### 4.4.2 Identification

To identify $\Theta$, I need to define a set of auxiliary models that provides a rich description of the data and captures the main features of the model. To achieve this goal, I use the following auxiliary models: (i) Linear Probability Models for choices (LPM), (ii) Mincer regressions for wages and criminal earnings,[28] (iii) variances from wages and criminal earnings, and (iv) variances and covariances from the residuals of the LPM. The estimated parameters from these auxiliary models are given in Appendix C.1.

The LPM have the following specification:

$$\begin{aligned} d_m(a) =& \beta_0^m + \beta_1^m \cdot a + \beta_2^m \cdot x_1(a) + \beta_3^m \cdot x_2(a) + \beta_4^m \cdot s(a) + \beta_5^m \cdot trg(a) + \beta_6^m \cdot mh(a) \\ &+ \beta_7^m \cdot nc(a) + \beta_8^m \cdot c + \beta_9^m \cdot (1 - d_3(a-1) - d_6(a-1)) + \beta_{10}^m \mathbb{1}_{(16 \le a \le 18)} \\ &+ \beta_{11}^m \mathbb{1}_{(18<a)} + \beta_{12}^m \mathbb{1}_{(16 \le a \le 18)} \cdot a + \beta_{13}^m \mathbb{1}_{(18<a)} \cdot a + \varepsilon_m(a), \end{aligned} \tag{4.19}$$

for $m = 1, 2, 3, 4$, and

$$\begin{aligned} d_m(a) =& \delta_0^m + \delta_1^m \cdot a + \delta_2^m \cdot x_1(a) + \delta_3^m \cdot x_2(a) + \delta_4^m \cdot s(a) + \delta_5^m \cdot trg(a) + \delta_6^m \cdot l(a) \\ &+ \delta_7^m \cdot mh(a) + \delta_8^m \cdot nc(a) + \delta_9^m \cdot c + \delta_{10}^m \cdot j(a) + \delta_{11}^m \cdot (1 - d_3(a-1) - d_6(a-1)) \\ &+ \delta_{12}^m \mathbb{1}_{(16 \le a \le 18)} + \delta_{13}^m \mathbb{1}_{(18<a)} + \delta_{14}^m \mathbb{1}_{(16 \le a \le 18)} \cdot a + \delta_{15}^m \mathbb{1}_{(18<a)} \cdot a + \varepsilon_m(a), \end{aligned} \tag{4.20}$$

for $m = 5, 6$.

[28] As discussed in Section 4.2, for the legal sector, human capital is a function of criminal records, school, training, experience, experience squared, cognitive and noncognitive skills, mental health, and the depreciation rate. Criminal capital is a function of criminal records, experience, experience squared, cognitive and noncognitive skills, mental health, and the depreciation rate.

The LPM helps to identify the parameters from the reward functions different from wages and illegal earnings. The second group of models identifies the parameters governing the human capital accumulation process. Note they are the same (ln) wage and illegal earnings equations from the model (equations 4.4 and 4.7). Some clarifications are important to make. There are two constants in the rewards for employment: $\gamma_{1,1}$ and $e_1(13)$. $e_1(13)$ is identified from the constant in the Mincer regression for observed (ln) wages, and $\gamma_{1,1}$ is normalized to zero. Having normalized $\gamma_{1,1}$, it is possible to identify $\gamma_{2,1}$ from the criminal rewards, $e_3(13)$ from school rewards, and $e_4(13)$ from home production rewards. Finally, $e_2(13)$ is identified from the Mincer regression for observed (ln) illegal earnings.

The same argument applies to identify the constants in the rewards from the "in detention" alternatives. $\gamma_{5,1}$ is normalized to zero making it possible to identify $e_6(13)$. Finally, $e_5(13)$ is identified from the regression of observed (ln) illegal earnings on a constant.

The third and fourth groups of moments help to identify $\Omega$, which I assume has the following structure:

$$\Omega = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{13} & \sigma_{15} & \sigma_{16} \\ \vdots & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} & \sigma_{26} \\ \vdots & \vdots & \sigma_3^2 & \sigma_{34} & \sigma_{35} & \sigma_{36} \\ \vdots & \vdots & \vdots & \sigma_4^2 & \sigma_{45} & \sigma_{46} \\ \vdots & \vdots & \vdots & \vdots & \sigma_5^2 & \sigma_{56} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \sigma_6^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{13} & 0 & 0 \\ \vdots & \sigma_2^2 & \sigma_{23} & \sigma_{24} & 0 & 0 \\ \vdots & \vdots & \sigma_3^2 & \sigma_{34} & 0 & 0 \\ \vdots & \vdots & \vdots & \sigma_4^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \sigma_5^2 & \sigma_{56} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \sigma_6^2 \end{bmatrix}.$$

Because alternatives $m = 1, .., 4$ and $m = 5, 6$ are not available at the same time, I assume the covariances between their random components are equal to zero (i.e. $\sigma_{m5} = 0$, $\sigma_{m6} = 0$, for $m = 1, 2, 3, 4$). Hence, I have 13 parameters to identify from $\Omega$.

To identify the parameters associated with $m = 5, 6$, I first normalize $\sigma_6^2 = 1$. $\sigma_5^2$ is identified form observed criminal earnings in detention. Given $\sigma_5^2$ and the normalization, I can identify $\sigma_{56}$.

Regarding the other elements in $\Omega$, I can identify $\sigma_1^2$ and $\sigma_2^2$ from the observed wages and illegal earnings, respectively. As discussed by Keane et al. (2011), once the parameters from the wage equation are identified, if there exists a variable, like experience ($x_1$), that enters the rewards for employment but not the rewards for school, I can identify $\sigma_{13}$. The same argument applies to $\sigma_{24}$. In this case, the exclusion restriction is criminal experience ($x_2$). By setting $\sigma_4^2 = 1$, I can identify the remaining elements of $\Omega$.[29]

### 4.4.3 Parameters Estimated Outside the Model

A set of parameters is estimated outside the model using PTD data. First, the detention probabilities ($\pi_m$) are the incarceration frequencies as a function of individuals' previous choices. Second, the probability of juvenile placement is given by the observed frequencies as a function of age. Third, I assume the distribution of $l(a + 1)$ for those first entering prison is equal

[29] By taking differences with respect to the first alternative, I end up with a system of five equations and five unknowns. By solving the system of equations, I can identify the remaining five elements from $\Omega$.

to the distribution of actual prison stay lengths from my sample. I estimate such a distribution via a duration model conditional on experience in the legal sector, criminal experience, years of school, years of training, age, and a dummy for having criminal records. Given the model estimates, I calculate the conditional probability of having a penalty of length $l$ at age $a+1$ and use those conditional probabilities in the model.

Fourth, I estimate the parameters for the law of motion of mental health and noncognitive skills outside the model. By doing so, I reduce the number of parameters to be estimated within the model and thus reduce the overall computational burden. There are two reasons I can do this. First, I observe these measures for everybody, in almost all periods, and the missing data is not correlated with observables. Therefore, there is no evidence of the existence of a selection process that could bias the estimates.[30] Second, given the timing assumption about the evolution of mental health and noncognitive skills, there is no endogeneity, and the OLS estimates for 4.12 and 4.13 should be consistent.[31] Finally, the distributions of $c$, $nc(13)$, $mh(13)$ are equal to the frequencies observed in the data, and the distribution of $\boldsymbol{\xi(a)}$ is obtained from the estimation of equations 4.12 and 4.13.

## 4.5 Results

In this section, I present the estimation results organized into two sections. The first section shows the goodness of fit of the model and provides a discussion of how well the model replicates the choice patterns observed in the data. In the second section, I comment on the estimates.

### 4.5.1 Goodness of Fit

Figures 4.2 to 4.7 compare the observed and simulated choices. Choices are expressed in regards to participation rates for each alternative. In general terms, the model does a good job of capturing the shape and the level for all choices, especially for crime (Figure 4.3), employment (Figure 4.2), training (Figure 4.7), and crime in detention (Figure 4.6).

The biggest discrepancies occur in school attendance (Figure 4.4) and home production (Figure 4.5) for ages 17 to 20. My estimates overpredict the decay in schooling and underpredict the increase in home production for this age range. This suggests that changes in the consumption value by age are not enough to explain why some people leave school earlier than others and more heterogeneity (e.g. discrete types) needs to be added to the model. However, my model does a job predicting participation in the long-run for these choices.

[30] In my panel, close to 1% of the observations for noncognitive skills were missing. For mental health, the proportion of missing observations was 19.2%.

[31] In particular, equations 4.12 and 4.13 define $mh(a)$ and $nc(a)$ as an autoregressive process that depends on the incarceration experience from the previous period. An endogeneity issue may arise if there is something correlated with the experience of incarceration that is also in the error term. Since I assume that $\boldsymbol{\xi(a)}$ is serially uncorrelated and independent of all past state variables, then no correlation with incarceration at $a-1$ is possible.

Figure 4.2: Percentage in employment by age

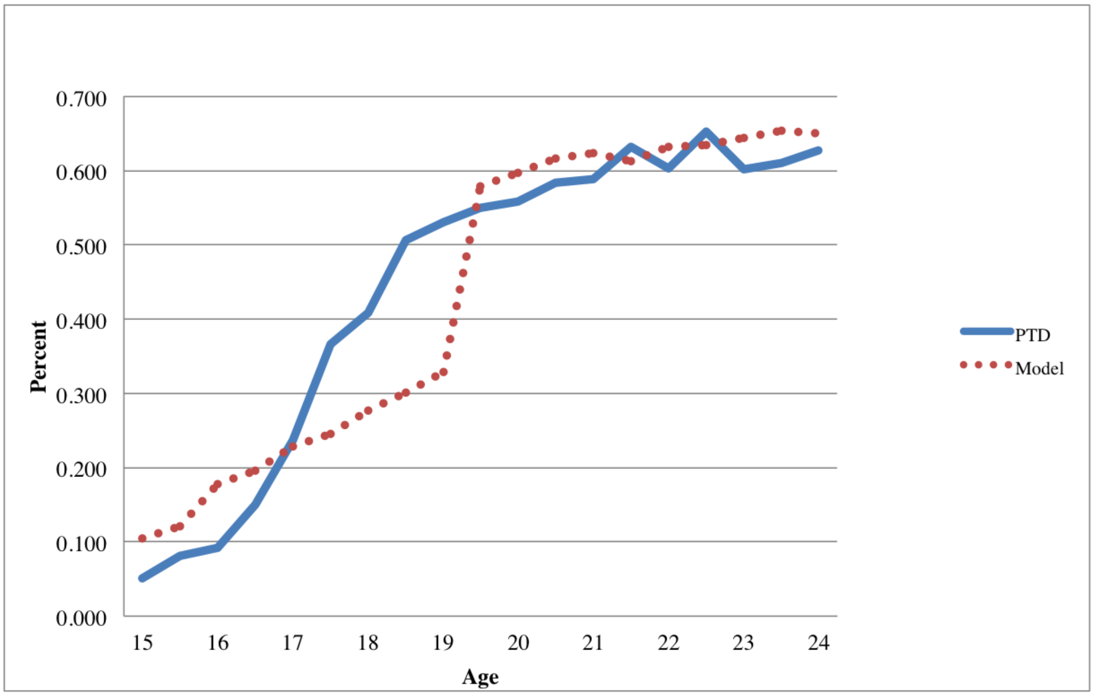



Figure 4.3: Percentage in crime by age

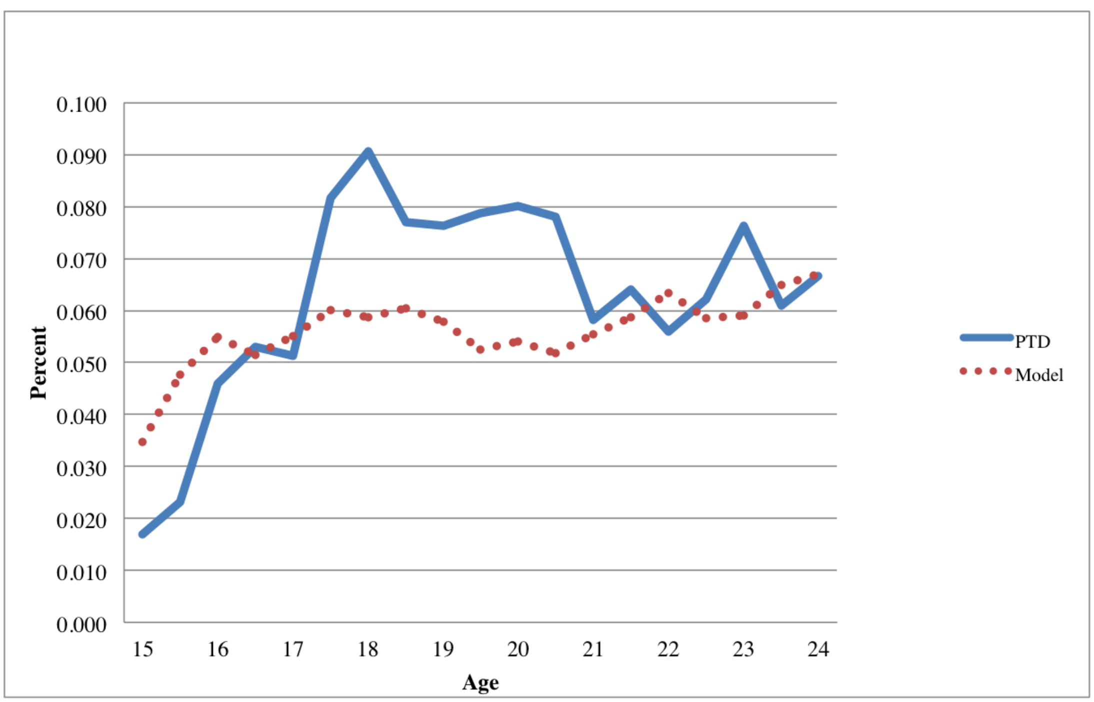


Figure 4.4: Percentage in school by age

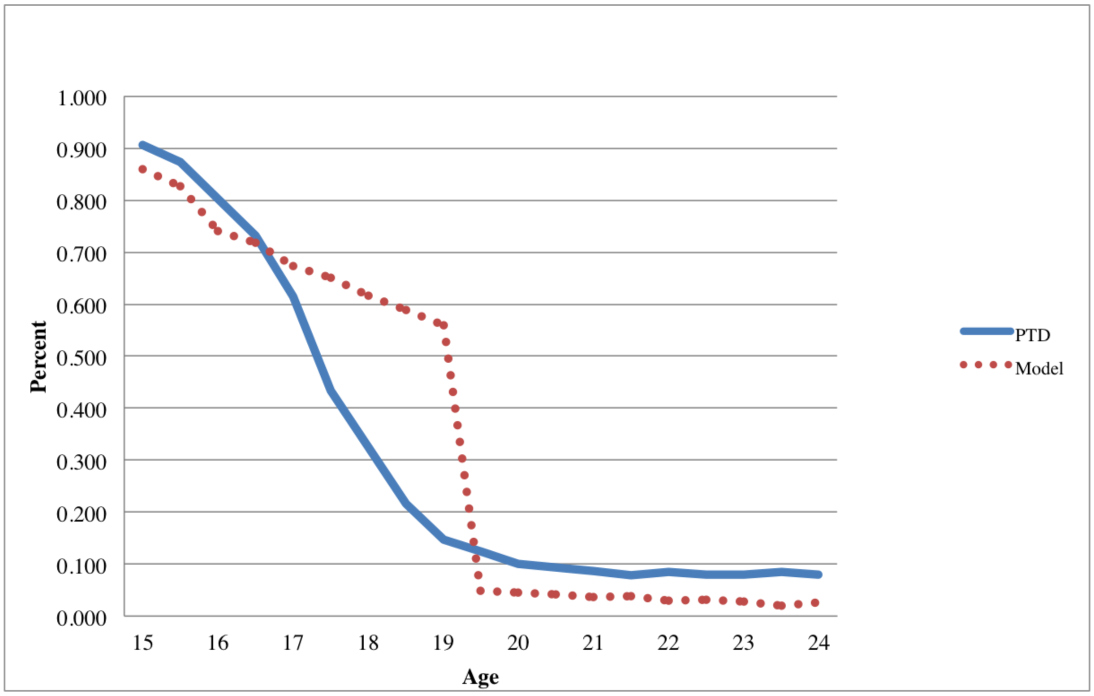



Figure 4.5: Percentage at home by age

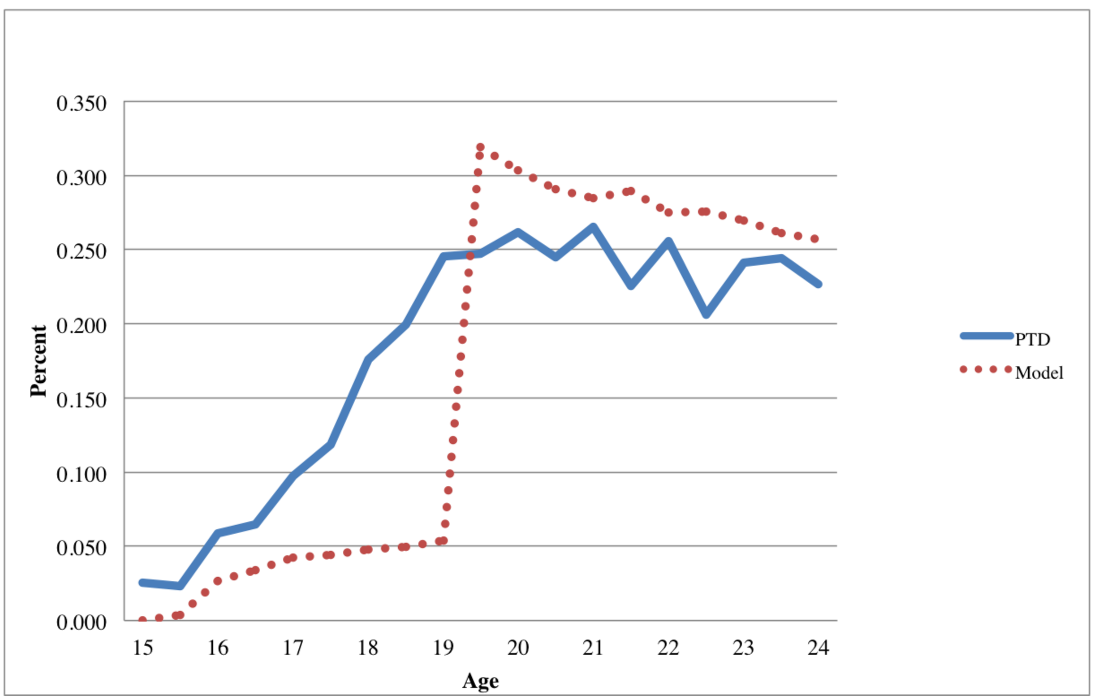


Figure 4.6: Percentage in crime in detention by age

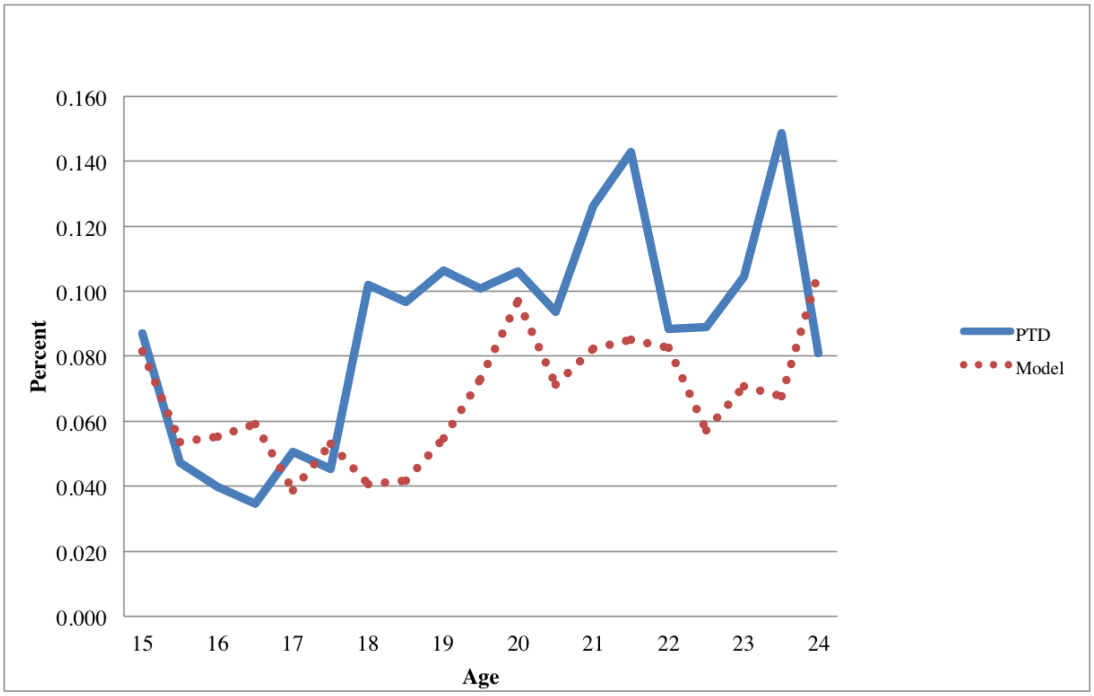
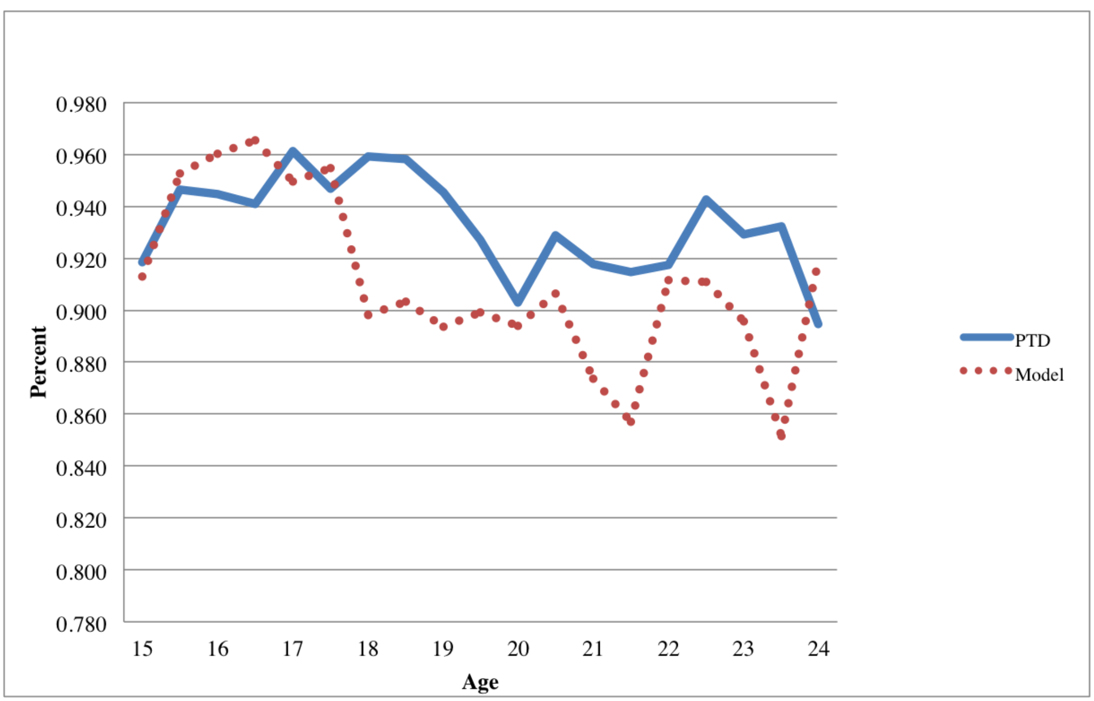



Figure 4.7: Percentage in training in detention by age



### 4.5.2 Parameter Estimates

Table 4.5 reports the parameter estimates from the human and criminal capital functions. According to my estimates, an additional semester of schooling increases wages in the legal sector by 6.6%, whereas an additional semester of training increases wages by 1.0% and this coefficient is not significant at conventional levels. An additional year of experience in the legal sector increases wages around 3.2%. The estimate for the quadratic term on experience is almost zero, which may be explained by the fact that wages are calculated only until individuals are 30 years old.

The impact of criminal records on wages is negative but small and nonsignificant. This finding is compatible with the fact that the majority of jobs from PTD respondents are low skilled and under-the-table jobs that are not provided by big employers. In these jobs, checks for criminal records are usually not requested by employers (Pager, 2003) and therefore ex-felons do not face different wages. Recently, some studies have provided evidence that former convicts face discrimination and adverse labour market conditions (Agan and Starr, 2017; Pager, 2003; Holzer et al., 2006). However, my results suggest that such conclusions do not extend to the type of legal jobs PTD participants find.

Cognitive skills and self-control have positive, but small returns. For example, an increase of one standard deviation in cognitive skills increases wages 2.3%. These findings are not surprising given that the majority of jobs for the people in the sample are low-skilled. Finally, there are positive returns to mental health. An increase of one standard deviation in this measure increases wages by 0.7%. However, the parameter is not precisely estimated.

Regarding criminal capital, note that returns to experience are higher than in the legal sector. The first year of criminal experience increases illegal earnings by 18%. Furthermore, there is evidence of decreasing returns to experience since the estimate for the squared term of experience is -0.005. Earnings in the criminal sector peak at approximately 7.5 years of continuous criminal experience, which implies that illegal earnings increase fast and for a short period.[32] Also, having a criminal record increases illegal earnings by 64.9% confirming the contribution of time behind bars to criminal capital (Bayer et al., 2009). However, this coefficient is not significant at conventional levels.

Cognitive ability has a positive contribution to criminal capital. According to my estimates, an increase of one standard deviation in this measure increases illegal earnings by approximately 22.3%. Not surprisingly, more intelligent offenders earn more money. In contrast, noncognitive ability has a negative contribution to criminal skills. A decrease of one standard deviation in noncognitive ability increases earnings by 2.4%, but the coefficient is not significant at conventional levels. One possible explanation consistent with such finding is that individuals with lower self-control take riskier opportunities and earn more money. Lastly, as in the legal sector, people with lower mental health earn less money.

Table 4.6 contains the estimates for the parameters of the rewards for school, home production, and training. Concerning school, it is important to highlight the rapid decrease in the net value of enrolment. A student who is 15 years old, independently of skills and mental health, has an estimated value of $1002.19, whereas for a 20-year-old student the estimated

[32]Loughran et al. (2013) report that going from low to a high level of experience would increase wages by more than 100%. My estimates suggest that an offender with criminal records would get an equivalent increase after working for 1.5 years in the criminal sector.

Table 4.5: Estimated Employment and Crime Parameters

| Variable | Employment | | Crime | |
|---|---|---|---|---|
| | Skill Functions | | | |
| Constant | 4.925*** | (0.0028) | 3.304*** | (3.3047) |
| Schooling | 0.066*** | (0.001) | … | |
| Trainning | 0.010 | (0.0076) | … | |
| Legal Experience | 0.032*** | (0.0016) | … | |
| Legal Experience Squared | 0.000 | (0.0004) | … | |
| Criminal Experience | … | | 0.18*** | (0.0069) |
| Criminal Experience Squared | … | | -0.005*** | (0.0009) |
| Ever in Detention | -0.001 | (0.0207) | 0.649 | (0.8385) |
| Cognitive Ability | 0.023*** | (0.0003) | 0.223*** | (0.0007) |
| Noncognitive Ability | 0.002 | (0.0187) | -0.024 | (0.1557) |
| Mental Health | -0.007 | (0.0207) | -0.030 | (0.1372) |
| Depreciation | -0.001 | (0.0026) | -0.001 | (0.5207) |
| Technology Shock§ | 0.308*** | (0.031) | 0.553*** | (0.1276) |
| | Nonpecuniary Benefits | | | |
| Constant | … | | 168.211*** | (9.6126) |
| Noncognitive Ability | … | | -4.994*** | (0.5021) |

Notes:
§ Estimated variance
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.6: Estimated School, Home and Training Parameters

| Variable | School | | Home | | Training | |
|---|---|---|---|---|---|---|
| Constant | 1111.999*** | (0.008) | 50.148*** | (0.0261) | 453.282*** | (0.0142) |
| Age Dummies: | | | | | | |
| Age 16-18 | -22.027*** | (8.4466) | … | | -0.009 | (0.8089) |
| Age 18.5 and over | -781.63*** | (14.6947) | 1199.302*** | (372.5299) | -54.8*** | (0.0772) |
| Age Slopes: | | | | | | |
| Age | -3.66*** | (0.0011) | 25.921*** | (0.0613) | 0.907*** | (0.0028) |
| Aged 16-18 | -2.948*** | (0.0013) | … | | 0.806*** | (0.0261) |
| Aged 18.5 and over | 4.694*** | (0.0019) | -26*** | (0.0008) | -0.892*** | (0.0024) |
| Reentry Cost | -41.465*** | (8.5939) | … | | -86.409*** | (0.0796) |
| Cognitive Ability | 2.175*** | (0.0004) | 0.885*** | (0.0002) | 0.415*** | (0.0003) |
| Noncognitive Ability | 6.359 | (8.3003) | 9.101*** | (0.0441) | 0.064** | (0.0323) |
| Mental Health | -4.034 | (8.2981) | -3.393*** | (0.0246) | 0.438*** | (0.0258) |
| Random Component§ | 0.949 | (4.3456) | 1.000 | … | 1.000 | … |

Notes:
§ Estimated variance
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

value is \$371.56.[33] This finding has important policy implications since additional incentives for school attendance should be given to individuals just at the end of compulsory schooling. Also, re-entry costs are low (\$41.46) and are unlikely to explain why people who leave school do not return. Note that more self-control increases the value of schooling, but the magnitude of the increase is small and nonsignificant.

Regarding home production, it is worth emphasizing that the net value of this alternative increases with age. For example, the value of this alternative, independently of skills and mental health, increases from \$828.34 when $i$ is 15 to \$1246.24 when $i$ is 20 years old.

To gain a better understanding of the overall effect of noncognitive skills on long-term choices, I use my estimates to simulate outcomes until age 30 given an increase of one standard deviation in noncognitive skills at age 13 (i.e. the first period). My results indicate that, by age 30, such an increase generates a decrease of 3.5 percentage points in criminal participation, a decrease of 3.0 percentage points in home production, and an increase of 1.5 percentage points in employment. Finally, by age 30, the population in detention decreases by 1.3 percentage points. These long-run effects are fundamentally coming from the reduction in illegal earnings associated with more self-control. These results confirm the negative effect of self-control on crime and its positive effect on employment and education found by Gottfredson and Hirschi (1990), Moffitt et al. (2011), and Heckman et al. (2006). Estimates from Chapter 3 also suggest an important effect of self-control on illegal earnings, but they are too noisy to derive further conclusions. In addition, the statistical models used in Chapter 3 do not account for all of the sources of dynamics incorporated in the structural model in this chapter. Therefore, in Chapter 3, I was not able to disentangle the long-run implications of changes in self-control on different choices.

## 4.6 Policy Experiments

In this section, I use the model's estimates to evaluate how individuals react to incentives in a dynamic context. I perform two policy experiments: school and wage subsidies. I focus on a school subsidy since the estimated returns for school are higher than for training. Also, a training subsidy would increase the net value of incarceration which makes crime more attractive.

For both policies, I simulate the choices up to age 30 and evaluate how criminal engagement changes over time. For the simulations, both policies are known and anticipated by the agents. An essential element of the discussion is the optimal timing of such policies. People in my sample are adolescents who are at a critical period of human capital accumulation and have high criminal engagement. It could be argued that "early" interventions may produce more significant long-term effects in decreasing crime and increasing employment. To test this prediction, I conduct two simulations for each policy: "early" and "late" policy interventions. For example, the "early" school subsidy is a transfer of \$300 per month to people who are in school between ages 16-18. The "late" school subsidy is a transfer of the same amount of money for people who are in school between ages 22-24. The "early" and "late" wage subsidies are also fixed transfers of \$300 per month for people who decide to work. The "early" transfer represents around 23% of the average potential monthly wage for individuals who decide to work at

[33] For a 15 years old student, this value comes from calculating $1111.99 - (3.66 \cdot 30)$

age 17 and around 20% for those who decide to work at age 23. Results for the wage subsidy are presented in Section 4.6.1 and for the school subsidy in Section 4.6.2. In Section 4.6.3, I present a back-of-the-envelope calculation to determine which one of these alternatives is the most effective crime-fighting strategy.

### 4.6.1 Wage Subsidy

Figures 4.8 to 4.11 show the simulated evolution of choices up to age 30 for a wage subsidy. These simulations show that behavioural responses from the "early" and "late" subsidy are different.

#### "Early" Wage Subsidy

The effects of the "early" subsidy on the short-run and long-run stocks of human and criminal capital are different. Compared to the no-policy levels, in the short-run (age 19), the intervention generates lower levels of human and criminal capital. In the long-run (age 30), the stock of human capital is lower, but the stock of criminal capital is almost the same. The stock of self-control is also affected by the policy.[34] Since lower levels of self-control are associated with more criminal engagement, this channel reinforces the effects on human and criminal capital.

Figure 4.8 shows the impact of the "early" intervention on employment. Between the ages 16-18, the increase in employment is significant, but it decreases sharply when the subsidy ends. The human capital level of marginal workers at age 19 is lower than the level when no policy is implemented.[35] This is because workers on the margin substitute school for employment. However, the estimated returns from working are half the returns from school. Since the marginal student drops out from school earlier (Figure 4.10), wages are lower, and employment is below the no-policy level.[36]

Regarding criminal participation (Figure 4.9), during the subsidy period, the "early" policy affects the opportunity cost of crime in two ways. First, it increases wages which discourage criminal participation and the accumulation of criminal capital. Second, it increases the incarceration cost since foregone earnings are higher, which also discourages criminal participation. In the short-run, criminal participation decreases without ambiguity and the level of criminal capital is lower.[37]

In the long-run, criminal participation goes back to its original level. The lower level of schooling generated by the policy (Figure 4.10) decreases the opportunity cost of crime in the long-run. Since criminal capital accumulates faster than human capital, after the subsidy ends marginal criminals come back to crime and compensate for the decline in criminal capital.[38]

[34] By age 30, the average level of self-control is 3.8% lower under the policy.

[35] At age 19, human capital for the marginal workers is 2.8% lower than the no-policy level. Also, the average human capital at age 19 for the "early" policy is 3% lower than the human capital when no policy is implemented.

[36] At age 19, for the "early" policy, the simulated average wages for participants of the labour market are 2.5% lower than the wages when no policy is implemented.

[37] According to my simulations, at age 19, the average simulated criminal capital is 7.3% lower than when no policy is implemented.

[38] To illustrate this point, Figure C.1 presents the evolution of human capital and criminal capital for two policy regimes: no policy and "early" wage subsidy for the group of marginal "permanent" offenders. A "permanent" offender is a person who is in the 90 percentile of criminal experience by age 30. That means people who have

Given the lower wages and the criminogenic effect of incarceration, illegal earnings rise fast causing marginal offenders to stay in crime. By age 30, criminal participation and average criminal capital are almost the same as without the policy.[39]

For home production, the effects are also different in the short- and long-run. During the subsidy period, the economic incentives to work are higher. People temporarily switch home production for employment, but working does not contribute as much to the creation of human capital. When the subsidy ends, human capital for the marginal workers is not high enough to create a permanent incentive to stay employed, and so many of them return to home production. In fact, by age 30 there is an increase of 3 percentage points in home production (Figure 4.11). In general, in the long-run, the "early" wage subsidy produces more high school dropouts, lower wages, and more people choosing to stay at home.

### "Late" Wage Subsidy

In general terms, the short- and long-run effects for the "late" policy are not contradictory. The intervention produces an increase in employment, a decrease in criminal participation, and a reduction in home production. Interestingly, the interplay between different choices affects the accumulation of human and criminal capital before, during, and after the subsidy.[40]

Before the subsidy period, there is an increase in school attendance and employment. There is also a decrease in criminal participation and home production. Since expected wages are higher due to the subsidy, there is an incentive to accumulate experience and years of schooling before the subsidy period to increase the chances of getting the subsidy. As a result, the policy encourages early investments in human capital because it is anticipated. This is a fundamental difference with the "early" subsidy.

More human capital increases the opportunity cost of crime since forgone earnings are higher. Also, investments in human capital stop the accumulation of criminal capital which reduces the rewards from criminal participation. Figure 4.9 shows that criminal participation decreases by 2.4 percentage points by age 30. In addition, more human capital increases the

---

accumulated at least 15 periods of criminal experience. I defined as marginal, the person who changes crime for employment during the policy period. For example, for the "early" wage subsidy, $i$ is said to be marginal if he commits a crime in all the periods when he is between 16 and 18 years old under no policy. When the policy is implemented the same person chooses employment for all the same periods (i.e. when he is between 16 and 18 years old).

Figure C.1 shows that under no policy, criminal capital always lies above human capital. When the policy is introduced criminal capital declines and human capital increases to the point where, during the subsidy period, criminal capital is below human capital. But when the subsidy period ends, criminal capital starts to increase faster than human capital. By age 20, criminal capital has completely overtaken human capital. In the long-run, the average criminal capital for "permanent" offenders is even higher than at the no policy level.

The situation is different from the "late" wage subsidy. Figure C.2 shows that criminal capital is below human capital before the subsidy period. By age 20, criminal capital intersects human capital and stays closer to it. As a result, crime is no longer the obvious choice for this group. The average criminal experience by age 30 is 23.8 periods under no policy, 17.5 periods for the "early" policy, and 13 periods for the "late" policy.

[39] In fact, by age 30, the average level of criminal capital with the policy is 0.7% lower than without the policy.

[40] For example, by age 19, before the policy is implemented, human capital is 4.5% higher, and criminal capital is 11.5% lower than at the no-policy level. By the end of the subsidy, age 24, human capital is 10.3% higher, and criminal capital is 15.3% lower. Finally, by age 30, human capital is 14.5% higher, and criminal capital is 14.7% lower. Also, self-control is higher in each of those ages reinforcing the role of human capital.

opportunity cost of home production making this alternative less attractive before the subsidy period. Figures 4.9 and 4.11 show the described changes before the subsidy period. During the subsidy period, employment increases sharply, and more human capital is accumulated which further discourages the accumulation of criminal capital.

When the policy ends, there is a decrease in employment, but the employment rate is still above the no-policy level. Higher levels of human capital make employment the preferred alternative for the majority. By age 30, employment is 13 percentage points higher than without the policy.

Figure 4.8: Percentage employment by age (wage subsidy).

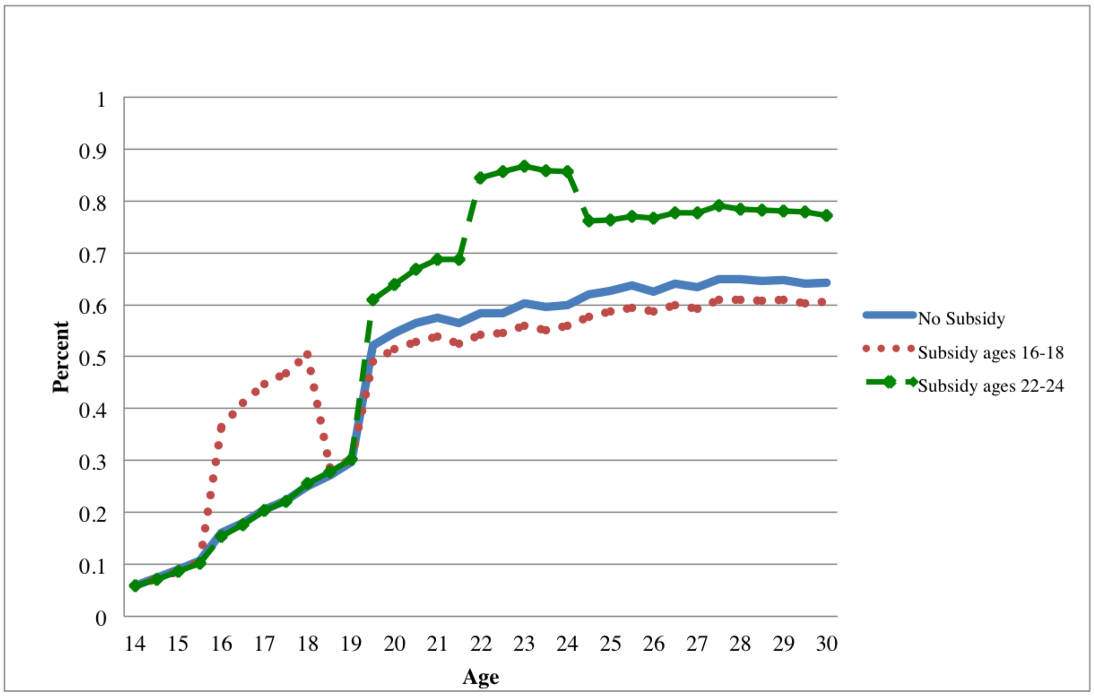


Figure 4.9: Percentage in crime by age (wage subsidy).

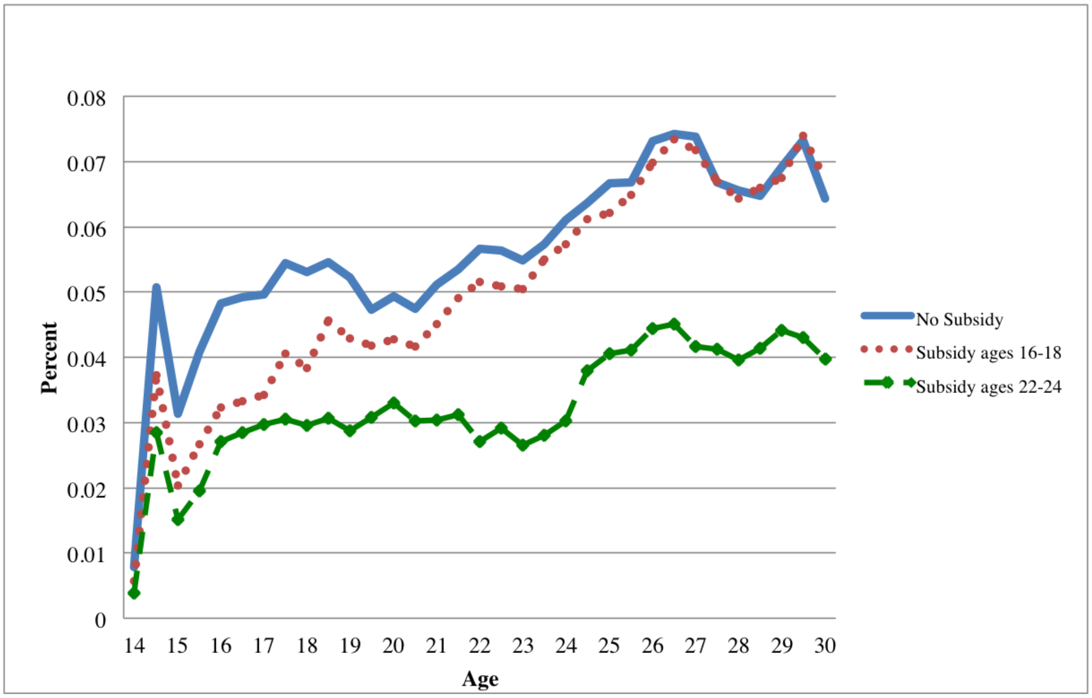



Figure 4.10: Percentage in school by age (wage subsidy).

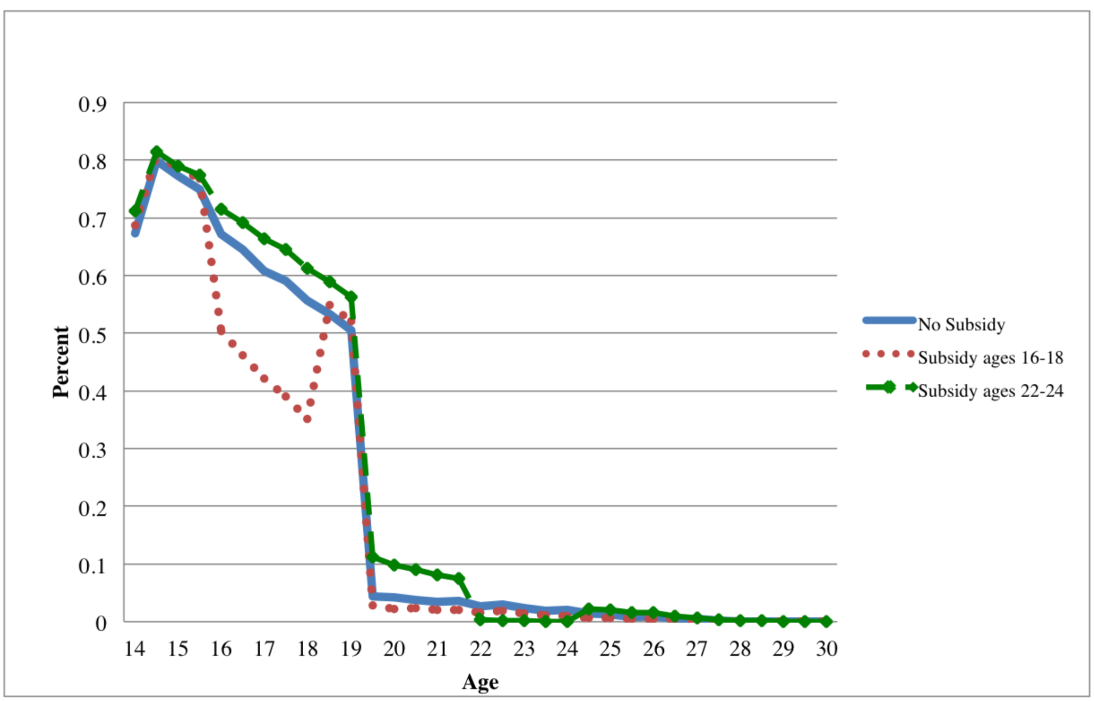
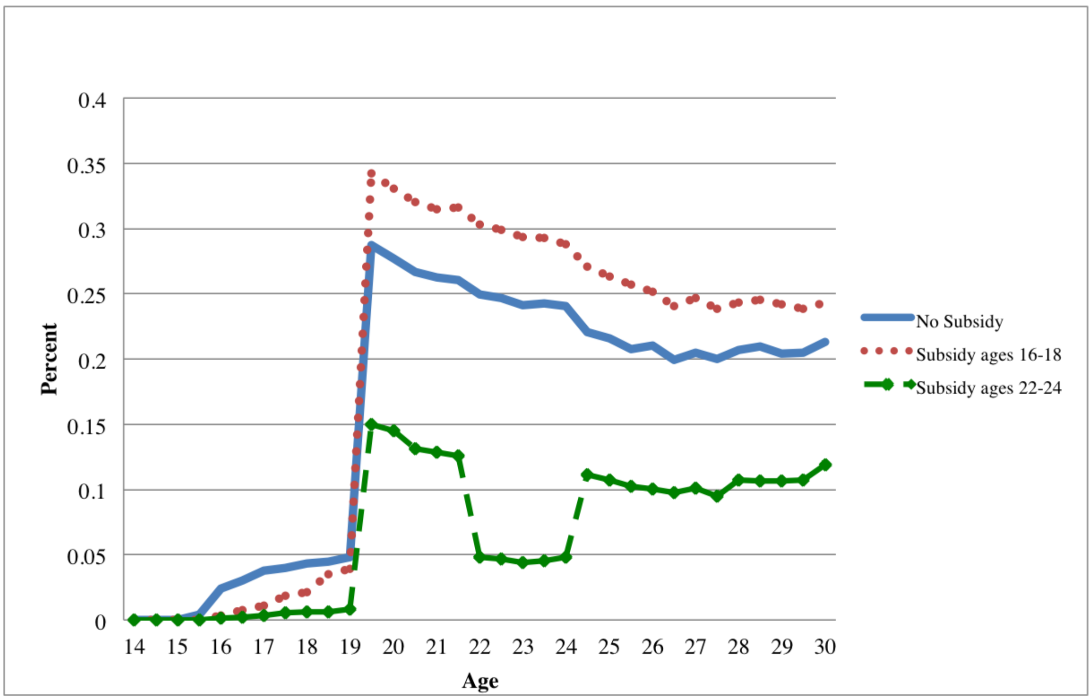


Figure 4.11: Percentage at home by age (wage subsidy).



### 4.6.2 School Subsidy

The "early" and the "late" school subsidy produce similar long-run effects over employment, crime and home production. However, the policy that creates more incentives for early investment, generates the biggest effects.

#### "Early" School Subsidy

Figures 4.12 to 4.15 show the simulated evolutions of crime, employment, school, and home production participation between 14 and 30 years old. In Figure 4.14 it is clear that there is an increase in school participation during the subsidy period. Thanks to the policy, students on the margin stay in school. Therefore, there is a decrease in employment. Part of the reduction is associated with the increase in the consumption value for school. Another part is associated with the expected increase in future wages. Because school has the most significant contribution to the evolution of human capital, by age 19, human capital is 5.7% higher than without the policy. Also, there is a decrease in criminal participation. The increase of the net value for school increases the opportunity cost of crime which prevents the further accumulation of criminal capital.[41] As a consequence, criminal engagement is lower than the no-policy level after the subsidy period. In the long-run (by age 30) the reduction in criminal participation is around 2 percentage points, and the decrease in criminal capital is close to 11%.

This reduction is bigger than the one found in Fella and Gallipoli (2014) using an equivalent policy.[42] Two reasons may explain the difference in results. First, Fella and Gallipoli (2014) do not allow for the endogenous evolution of criminal capital, and individual heterogeneity is fixed in their model. Therefore, their sources of dynamics for criminal engagement are more restricted than the ones presented here. In particular, their model does not account for the fact that any reductions in crime limit criminal experience, which further discourages future

[41] Figures C.3-C.8 present additional simulations to illustrate how the accumulation of human capital and criminal capital interact over time for the group of "permanent offenders" I defined before. I choose this group simply because they are the most likely to come back to crime when the policy period ends. I investigate four new policy regimes for people 16-18 years old: (i) mandatory employment, (ii) mandatory employment without incarceration, (iii) mandatory schooling, and (iv) mandatory schooling without incarceration. My results indicate that, in some cases, the rapid growth of criminal capital quickly overtakes human capital and erase the temporary reduction in criminal engagement due to the policy. However, when the human capital created by the policy is high enough, individuals face an increasing opportunity cost for criminal engagement which prevents the further accumulation of criminal capital.

Figure C.3 shows that for the early wage subsidy, criminal capital is higher than human capital soon after the subsidy period ends. The same pattern is observed under mandatory employment (Figure C.4) and mandatory employment without incarceration (Figure C.5). In both cases, when the policy period ends, "permanent" offenders start accumulating criminal capital again. That accumulation is slower without incarceration, but eventually criminal capital overtakes human capital.

For the case of mandatory schooling (Figure C.7) and mandatory schooling without incarceration (Figure C.8), the conclusion is different. In both cases, criminal capital never exceeds human capital creating a significant reduction in the long-run criminal engagement of "permanent" offenders. For the "early" school subsidy (Figure C.6), criminal capital exceeds human capital after the subsidy period ends and just for a short period. Because of the existence of decreasing returns to criminal experience and the continuous accumulation of experience in the legal sector, criminal capital lies below human capital by age 30. In these school policies, high levels of human capital prevent criminal capital to overtake it.

[42] Fella and Gallipoli (2014) find a reduction in the crime rate of 8% while my simulations imply a reduction of 30% of criminal participation by age 30.

crime and creates a feedback that further encourages investments in human capital. Second, my model estimates come from a population of disadvantaged youth who are at the bottom of the ability distribution, whereas results from Fella and Gallipoli (2014) come from the overall population.[43]

### "Late" School Subsidy

Figure 4.12 shows that employment is almost the same, up to age 22, between "late" school subsidy and no policy. The same happens with school attendance (Figure 4.14) and criminal participation (Figure 4.13). During the subsidy period, the "late" policy increases school participation and decreases employment, home production, and crime. The reasons are the same as discussed for the "early" policy: an increase in the net value of school increases the opportunity cost of crime and home production. People with low wages prefer to substitute employment for school since the returns for the latter are higher.

In addition, the "late" intervention does not provide enough incentives for people to make investments in education. Most people in my sample have finished with school by age 20. Older people have the lowest valuation for school and get the highest wages and illegal earnings. When the subsidy period ends, human capital is higher and criminal capital is lower than at the no-policy level. However, human capital is around 5% lower and criminal capital is around 13% higher than for the "early" subsidy. These differences explain why the "early" intervention produces bigger effects by age 30.

Offenders in their early 20's already have accumulated a large amount of criminal capital. As a result, this policy does not provide enough incentives to generate crime reductions in the long-run as big as for the "early" school subsidy. The incentives to make investments in education arrive too late for a population of people who already have criminal records and low human capital. As a result, the long-run decrease in criminal participation, compared to the no-policy level, is only about 0.5 percentage points.

In any case, both types of school subsidy produce long-run effects on crime decisions. This conclusion is different than the one reached by Mancino et al. (2016) for enrollment policies. According to their results, policies based on temporary interventions will have only small effects on criminal behaviour many years after the policy. Three reasons may explain why our conclusions differ. First, my definition of criminal engagement is different.[44] Second, the endogenous evolution of illegal earnings is not affecting criminal decisions in Mancino et al. (2016). Third, I consider additional sources of dynamics for the evolution of human and criminal capital that impact crime in the long-run, like the effects of incarceration and the returns to criminal experience.

The fundamental lesson from these experiments is that policies for youth should encourage investments in education because school has the most significant contribution to human capital. Any policy that reduces the incentives to attend school before age 19 has a negative effect on the evolution of human capital, and it may increase crime and home production in the long-run.

[43] Fella and Gallipoli (2014) also find bigger effects on crime reduction for people at the lower end of the ability distribution.

[44] I focus on income-generating crimes while, for the simulations, Mancino et al. (2016) use a definition that combines all types of crime.

Figure 4.12: Percentage employment by age (school subsidy)

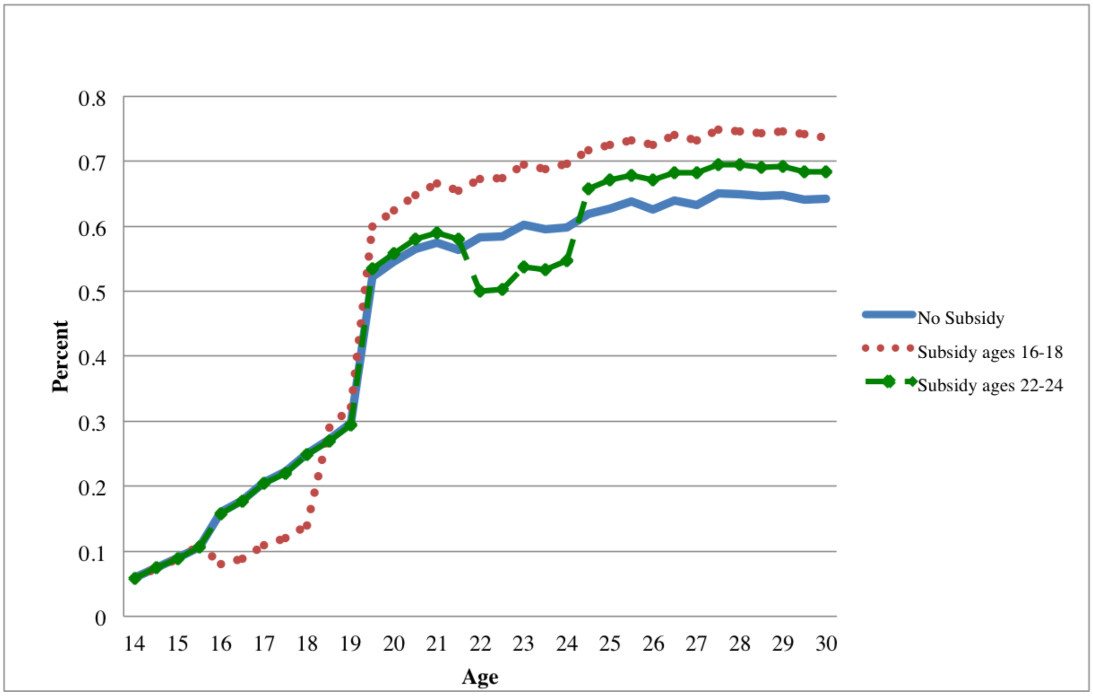



Figure 4.13: Percentage in crime by age (school subsidy)

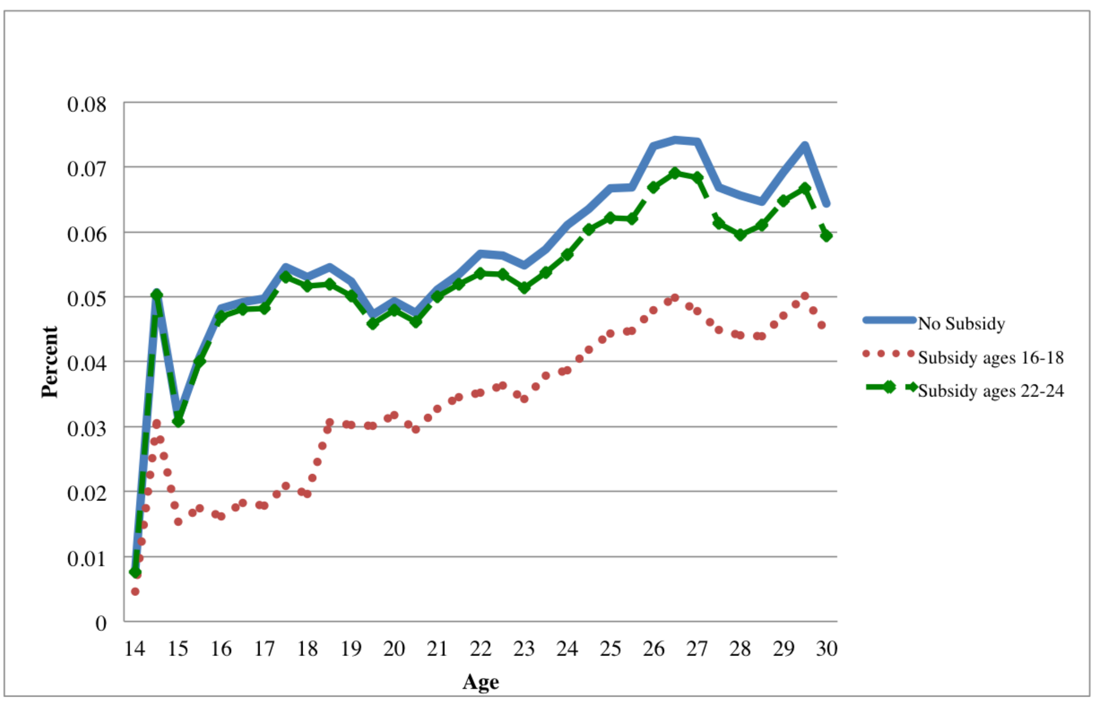


Figure 4.14: Percentage in school by age (school subsidy)

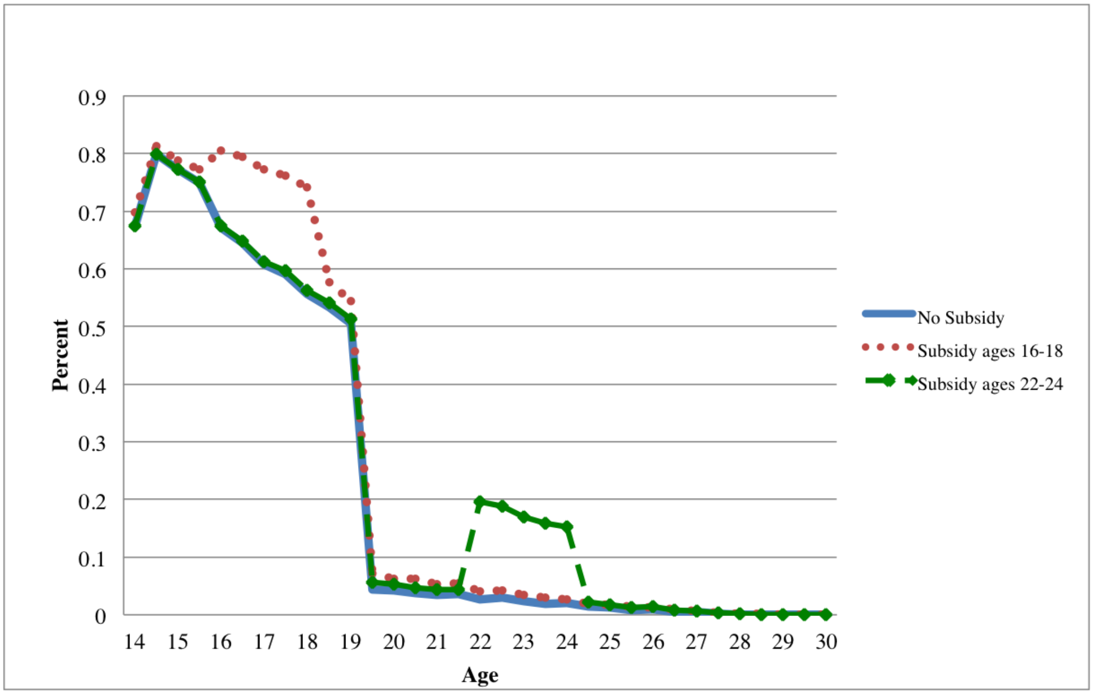
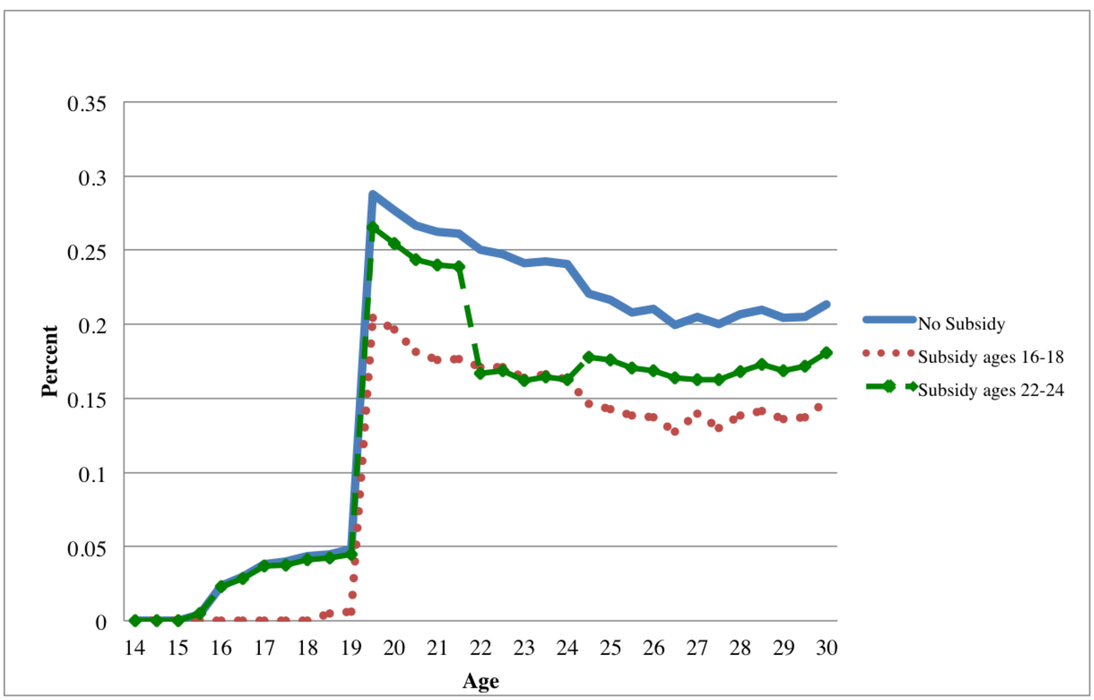



Figure 4.15: Percentage at home by age (school subsidy)



### 4.6.3 Cost Comparison: A Back-of-the-envelope Calculation

Table 4.7 presents the cost for all four simulated policies. The cost for each policy is calculated as the total number of participants during the subsidy period,[45] from a simulated sample of ten thousand individuals, times the amount of the transfer (which is $300 in all cases). There are several ways to compare the cost-effectiveness between policies. Here, I choose two: (i) the cost of reducing criminal participation by one percentage point compared to the no-policy level at age 30 and (ii) the cost of reducing the population in detention by one percentage point compared to the no-policy level at age 30.

Table 4.7: Simulated Policy Costs

| | Timing | |
|---|---|---|
| Policy | "Early" | "Late" |
| School | 11.6 | 3.1 |
| Wage | 6.5 | 12.8 |

Note: Cost expressed in millions of dollars for ten thousand individuals.

The “early” wage subsidy causes a long-run (i.e. at age 30) increase in criminal participation of 0.28 percentage points compared to the no-policy level, and therefore is not subject to comparison. The “late” wage subsidy causes a long-run decrease of 2.46 percentage points compared to the no-policy level.[46] That means that decreasing criminal participation by one percentage point costs approximately 5.2 million dollars with the “late” wage subsidy.[47] Regarding the school subsidy, a one percentage point decrease for the “early” policy costs 5.8 million dollars, while for the “late” policy it costs 6.2 million dollars.[48] According to this metric, the most cost-effective crime-fighting strategy is the “late” wage subsidy.

In regards to the population in detention at age 30, all policies, except for the “early” wage subsidy, generate reductions with respect to the no-policy level. Those reductions are 0.9 percentage points for the “early” school subsidy, 0.34 percentage points for the “late” school subsidy, and 1.15 percentage points for the “late” wage subsidy. Therefore, the cost of reducing the population in detention by one percentage point is 12.8 million dollars for the “early” school subsidy, 9.11 million dollars for the “late” school subsidy, and 11.3 million dollars for the “late” wage subsidy. Even though the “late” school subsidy is the policy with the smallest effect, it is also the least expensive and therefore the most cost-effective strategy to reduce the population of inmates. The reason behind this is the effectiveness of the policy among the participants.

[45]The total number of participants is calculated as the total number of participants in each period times the number of subsidy periods.

[46]Under no-policy, the participation rate in crime, at age 30, is 6.44%. Under the “late” wage subsidy, the participation rate in crime, at age 30, is 3.98%. Therefore, the difference in participation rate between the two policies is 2.46 percentage points, i.e. $\Delta$Crime=2.46.

[47]The number comes from Total Policy Cost / $\Delta$Crime

[48]According to my simulations, the “early” school subsidy causes a decrease of 2 percentage points in criminal participation by age 30, while the “late” subsidy causes a decrease of 0.5 percentage points.

For example, for the "late" wage subsidy, 93.3% of the former participants are free by age 30 whereas, for the "late" school subsidy, that proportion is 94.1%. This fact suggests that "late" school interventions have the potential of changing the pathways for a greater number of program participants.

## 4.7 Conclusion

In this chapter, I analyze the life-cycle choices of juvenile offenders. To study the evolution of choices, using a new and rich dataset of juvenile offenders, I estimated a dynamic model of employment, crime, and education where people face the possibility of incarceration and its consequences. This chapter provides strong evidence that a model with endogenous human and criminal capital that accounts for the dynamic evolution of personal capabilities does a good job explaining the life-cycle choices of juvenile offenders.

According to my estimates, criminal capital accumulates faster than human capital, mainly because of returns to experience and incarceration records. In the legal sector, the evolution of human capital is slower but I did not find evidence of decreasing returns to experience. Also, the most significant contribution to skills is made by the school. However, my estimates suggest that the net value of school attendance decays after age 16, making it difficult for this population to even graduate from high school.

My results suggest a rather complicated response from individuals to incarceration. First, incarceration increases criminal capital. Second, the return to training is low. Third, incarceration influences the evolution of personal capabilities. Finally, incarceration stops the process of human capital formation.

Given the agreement of a human capital-based theory of crime with the data and the complicated responses associated with incarceration, it seems reasonable to explore alternative deterrent mechanisms based on incentives to increase human capital. I discuss two policies: school and wage subsidies. Using the estimates from the model, I simulate the life-cycle effects of each intervention on schooling, crime, and employment. The results show that both policies can be used as effective crime-fighting strategies, but the age of the intervention is critical for producing long-run effects.

For the school subsidy, I find that the "early" intervention produces the most significant long-run effects on employment and crime reduction. For the wage subsidy, my results indicate the "early" intervention produces only a short-run effect on employment and crime reduction. After the subsidy period ends, marginal workers go back to crime and the effects of the policy disappear over time.

Because the model developed in this chapter accounts for the dynamic interaction between human and criminal capital over the life-cycle, it offers a framework to evaluate the age of optimal interventions.

# Bibliography


Agan, A. and S. Starr (2017, 08). Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment. *The Quarterly Journal of Economics 133*(1), 191–235.

Agency for Healthcare Research and Quality (2014). Mental health: Research findings. https://www.ahrq.gov/sites/default/files/publications/files/mentalhth.pdf.

Aizer, A. and J. Doyle (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *Quarterly Journal of Economics 130*(2), 759–803.

Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz (2011). Personality psychology and economics. In *Handbook of the Economics of Education*, Volume 4, pp. 1–181. Elsevier.

Anwar, S. and T. A. Loughran (2011). Testing a bayesian learning theory of deterrence among serious juvenile offenders. *Criminology 49*(3), 667–698.

Baćak, V., L. H. Andersen, and J. Schnittker (2019, September). The effect of timing of incarceration on mental health: Evidence from a natural experiment. *Social Forces 98*(1), 303–328.

Barnert, E. S., R. Perry, and R. E. Morris (2016). Juvenile incarceration and health. *Academic Pediatrics 16*(2), 99–109.

Barnes, J., K. M. Beaver, and J. M. Miller (2010). Estimating the effect of gang membership on nonviolent and violent delinquency: A counterfactual analysis. *Aggressive behavior 36*(6), 437–451.

Bayer, P., R. Hjalmarsson, and D. Pozen (2009). Building criminal capital behind bars: Peer effects in juvenile corrections. *The Quarterly Journal of Economics 124*(1), 105–147.

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2016). Incarceration, recidivism and employment. Working Paper.

Binswanger, I. A., M. F. Stern, R. A. Deyo, P. J. Heagerty, A. Cheadle, J. G. Elmore, and T. D. Koepsell (2007). Release from prison — a high risk of death for former inmates. *The New England Journal of Medicine 356*(2), 157–165.

Black, D. A. and J. A. Smith (2004). How robust is the evidence on the effects of college quality?: Evidence from matching. *Journal of econometrics 121*(1), 99;124;–124.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. t. Weel (2008). The economics and psychology of personality traits. *The Journal of Human Resources 43*(4), 972–1059.

Brand, J. E. and Y. Xie (2010). Who benefits most from college?: Evidence for negative selection in heterogeneous economic returns to higher education. *American Sociological Review 75*(2), 273,302.

Calcaterra, S. L., B. Beaty, S. R. Mueller, S.-J. Min, and I. A. Binswanger (2014). The association between social stressors and drug use/hazardous drinking among former prison inmates. *Journal of Substance Abuse Treatment 47*(1), 41 – 49.

Case, A. C. and L. F. Katz (1991, May). The company you keep: The effects of family and neighborhood on disadvantaged youths. Working Paper.

Caspi, A., T. E. Moffitt, P. A. Silva, M. Stouthamer-Loeber, R. F. Krueger, and P. S. Schmutte (1994). Are some people crime-prone? replications of the personality-crime relationship across countries, genders, races, and methods. *Criminology 32*(2), 163–196.

Cauffman, E., L. Steinberg, and A. R. Piquero (2005). Psychological, neuropsychological and physiological correlates of serious antisocial behavior in adolescence: The role of self-control. *Criminology 43*(1), 133–176.

Coker, K. L., P. H. Smith, A. Westphal, H. V. Zonana, and S. A. McKee (2014). Crime and psychiatric disorders among youth in the us population: An analysis of the national comorbidity survey–adolescent supplement. *Journal of the American Academy of Child and Adolescent Psychiatry 53*(8), 888–898.e2.

Cook, M. N., J. Peterson, and C. Sheldon (2009). Adolescent depression: an update and guide to clinical decision making. *Psychiatry (Edgmont (Pa. : Township)) 6*(9), 17.

Currie, J. and E. Tekin (2012). Understanding the cycle childhood maltreatment and future crime. *Journal of Human Resources 47*(2), 509–549.

Derogatis, L. R. and N. Melisaratos (1983). The brief symptom inventory: an introductory report. *Psychological Medicine 13*(3), 595–605.

Dick, B. and B. J. Ferguson (2015). Health for the world's adolescents: a second chance in the second decade. *Journal of Adolescent Health 56*(1), 3–6.

Donohue, J. J. and P. Siegelman (1998). Allocating resources among prisons and social programs in the battle against crime. *The Journal of Legal Studies 27*(1), 1–43.

Duckworth, A. L. and M. E. P. Seligman (2017, September). The science and practice of self-control. *Perspectives on Psychological Science 12*(5), 715–718.

Eriksson, K. H., R. Hjalmarsson, M. J. Lindquist, A. Sandberg, I. för social forskning (SOFI), I. för internationell ekonomi, S. universitet, and S. fakulteten (2016). The importance of family background and neighborhood effects as determinants of crime. *Journal of Population Economics 29*(1), 219–262.

Ettner, S. L., R. G. Frank, and R. C. Kessler (1997). The impact of psychiatric disorders on labor market outcomes. *Industrial and Labor Relations Review 51*(1), 64–81.

Fagan, J. (1996). The comparative advantage of juvenile versus criminal court sanctions on recidivism among adolescent felony offenders. *Law and Policy 18*(1-2), 77–114.

Fagan, J. and A. Kupchik (2011). Juvenile incarceration and the pains of imprisonment. *Duke FL & Soc. Change 3*, 29.

Farrell, M. and J. Marsden (2008). Acute risk of drug-related death among newly released prisoners in england and wales. *Addiction 103*(2), 251–255.

Fazel, S. and J. Danesh (2002). Serious mental disorder in 23000 prisoners: a systematic review of 62 surveys. *Lancet (London, England) 359*(9306), 545.

Fella, G. and G. Gallipoli (2014). Education and crime over the life cycle. *Review of Economic Studies 81*(4), 1484–1517.

Fletcher, J. M. (2008). Adolescent depression: diagnosis, treatment, and educational attainment. *Health economics 17*(11), 1215–1235.

Fletcher, J. M. (2010). Adolescent depression and educational attainment: results using sibling fixed effects. *Health economics 19*(7), 855–871.

Fletcher, J. M. (2012). Adolescent depression and adult labor market outcomes. Working Paper.

Frank, R. and T. G. McGuire (2010, April). Mental health treatment and criminal justice outcomes. Working Paper.

Fridell, M., M. Hesse, M. M. Jæger, E. Kühlhorn, I. för psykologi, L. University, D. of Psychology, and L. universitet (2008). Antisocial personality disorder as a predictor of criminal behaviour in a longitudinal study of a cohort of abusers of several classes of drugs: Relation to type of substance and type of crime. *Addictive Behaviors 33*(6), 799–811.

Gottfredson, M. R. and T. Hirschi (1990). *A general theory of crime*. Stanford University Press.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics 8*, S85–S118.

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

Haney, C. (2017). "madness" and penal confinement: Some observations on mental illness and prison pain. *Punishment and Society 19*(3), 310–326.

Harding, D. J., J. D. Morenoff, A. P. Nguyen, and S. D. Bushway (2018). Imprisonment and labor market outcomes: Evidence from a natural experiment. *American Journal of Sociology 124*(1), 49–110.

Hazel, N. (2008). Cross-national comparison of youth justice. Technical report, Youth Justice Board for England and Wales.

Heckman, J. J., H. Ichimura, and P. E. Todd (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies 64*(4), 605–654.

Heckman, J. J. and R. Robb (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics 30*(1), 239–267.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

Hjalmarsson, R. (2008). Crime and expected punishment: Changes in perceptions at the age of criminal majority. *American Law and Economics Review 11*(1), 209–248.

Hjalmarsson, R. and A. Bindler (2017). Prisons, recidivism and the age-crime profile. *Economics Letters 152*, 46–49.

Hoeve, M., L. S. McReynolds, and G. A. Wasserman (2013). The influence of adolescent psychiatric disorder on young adult recidivism. *Criminal justice and behavior 40*(12), 1368–1382.

Holzer, H. J., S. Raphael, and M. A. Stoll (2006). Perceived criminality, criminal background checks, and the racial hiring practices of employers. *The Journal of Law and Economics 49*(2), 451–480.

Imai, S. and K. Krishna (2004). Employment, deterrence, and crime in a dynamic model. *International Economic Review 45*(3), 845–872.

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

Johnson, R. C. and S. Raphael (2009). The effects of male incarceration dynamics on acquired immune deficiency syndrome infection rates among african american women and men. *The Journal of Law and Economics 52*(2), 251–293.

Keane, M. P., P. E. Todd, and K. I. Wolpin (2011). The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. In *Handbook of labor economics*, Volume 4, pp. 331–461. Elsevier.

Kehl, D. L. and S. A. Kessler (2017). Algorithms in the criminal justice system: Assessing the use of risk assessments in sentencing. Responsive Communities Initiative, Berkman Klein Center for Internet and Society, Harvard Law School.

Knapp, M., V. Ardino, N. Brimblecombe, S. Evans-Lacko, V. Lemmi, D. King, T. Snell, S. Murguia, H. Mbeah-Bankas, S. Crane, et al. (2016). Youth mental health: new economic evidence. Technical report, Personal Social Services Research Unit, London School of Economics and Political Science.

Knox, P. (2017). The risks and rewards of risk assessments. https://www.ncsc.org/microsites/trends/home/Monthly-Trends-Articles/2017/The-Risks-and-Rewards-of-Risk-Assessments.aspx.

Lamb, H. R. and L. E. Weinberger (1998). Persons with severe mental illness in jails and prisons: a review. *Journal of Clinical Forensic Medicine 5*(4), 213–213.

Land, K. C. and D. S. Nagin (1996). Micro-models of criminal careers: A synthesis of the criminal careers and life course approaches via semiparametric mixed poisson regression models, with empirical applications. *Journal of Quantitative Criminology 12*(2), 163–191.

Langan, P. A. and D. J. Levin (2002). Recidivism of prisoners released in 1994. Special Report NCJ 193427, Bureau of Justice Statistics.

Laub, J. H. and R. J. Sampson (1998). *The long-term reach of adolescent competence: Socioeconomic achievement in the lives of disadvantaged men*. University of Chicago Press Chicago, IL.

Layard, R. (2017). The economics of mental health. *IZA World of Labor 321*, 1–10.

Lechner, M. (2002). Program heterogeneity and propensity score matching: An application to the evaluation of active labor market policies. *The Review of Economics and Statistics 84*(2), 205–220.

Lerner, D. and R. M. Henke (2008). What does research tell us about depression, job performance, and work productivity? *Journal of Occupational and Environmental Medicine 50*(4), 401–410.

Leung, S. F. (1994). An economic analysis of the age-crime profile. *Journal of Economic Dynamics and Control 18*(2), 481–497.

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

Lochner, L. (2007). Individual perceptions of the criminal justice system. *The American Economic Review 97*(1), 444–460.

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

Loughran, T. A., H. Nguyen, A. R. Piquero, and J. Fagan (2013). The returns to criminal capital. *American Sociological Review 78*(6), 925–948.

Mancino, M. A. (2018). A search model of early employment careers and youth crime. Working Paper.

Mancino, M. A., S. Navarro, and D. A. Rivers (2016). Separating state dependence, experience, and heterogeneity in a model of youth crime and education. *Economics of Education Review 54*, 274–305.

Massoglia, M. (2008). Incarceration as exposure: The prison, infectious disease, and other stress-related illnesses. *Journal of Health and Social Behavior 49*(1), 56–71.

Massoglia, M. and W. A. Pridemore (2015). Incarceration and health. *Annual Review of Sociology 41*(1), 291–310.

Massoglia, M., B. Remster, and R. D. King (2011). Stigma or separation? understanding the incarceration-divorce relationship. *Social Forces 90*(1), 133–155.

McCarthy, B. and J. Hagan (2001). When crime pays: Capital, competence, and criminal success. *Social Forces 79*(3), 1035–1060.

Merlo, A. and K. I. Wolpin (2015). The transition from school to jail: Youth crime and high school completion among black males. *European Economic Review 79*, 234–251.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy 66*(4), 281–302.

Mocan, H. N., S. C. Billups, and J. Overland (2005). A dynamic model of differential human capital and criminal activity. *Economica 72*(288), 655,681.

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

Moffitt, T. E., L. Arseneault, D. Belsky, N. Dickson, R. J. Hancox, H. Harrington, R. Houts, R. Poulton, B. W. Roberts, S. Ross, M. R. Sears, W. M. Thomson, and A. Caspi (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences 108*(7), 2693–2698.

Morgan, S. L. (2001). Counterfactuals, causal effect heterogeneity, and the catholic school effect on learning. *Sociology of Education 74*(4), 341–374.

Mueller-Smith, M. (2015). The criminal and labor market impacts of incarceration. Working Paper.

Mulvey, E. P., C. A. Schubert, and H. L. Chung (2007). Service use after court involvement in a sample of serious adolescent offenders. *Children and Youth Services Review 29*(4), 518–544.

Mulvey, E. P., C. A. Schubert, and A. Piquero (2014). Pathways to desistance: Final technical report. Technical report, National Institute of Justice.

Nagin, D. and R. Paternoster (2000). Population heterogeneity and state dependence: State of the evidence and directions for future research. *Journal of Quantitative Criminology 16*(2), 117–144.

Nagin, D. S. and R. Paternoster (1993). Enduring individual differences and rational choice theories of crime. *Law and Society Review 27*(3), 467–496.

Nagin, D. S. and R. Paternoster (1994). Personal capital and social control: The deterrence implications of a theory of individual differences in criminal offending. *Criminology 32*(4), 581–606.

Ng, I. Y., X. Shen, H. Sim, R. C. Sarri, E. Stoffregen, and J. J. Shook (2011). Incarcerating juveniles in adult prisons as a factor in depression. *Criminal Behaviour and Mental Health 21*(1), 21–34.

Nguyen, H., T. A. Loughran, R. Paternoster, J. Fagan, and A. R. Piquero (2017). Institutional placement and illegal earnings: Examining the crime school hypothesis. *Journal of Quantitative Criminology 33*(2), 207–235.

Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology 108*(5), 937–975.

Peng, L., C. D. Meyerhoefer, and S. H. Zuvekas (2016). The short-term effect of depressive symptoms on labor market outcomes. *Health Economics 25*(10), 1223–1238.

Pezzin, L. E. (2004). Effects of family background on crime participation and criminal earnings: An empirical analysis of siblings. *Estudos Econômicos (São Paulo) 34*(3), 487–514.

Piquero, A. R., E. Cauffman, and L. Steinberg (2005). Psychological, neuropsychological and physiological correlates of serious antisocial behavior in adolescence: The role of self-control. *Criminology 43*(1), 133–176.

Porter, L. C. (2014). Incarceration and post-release health behavior. *Journal of Health and Social Behavior 55*(2), 234–249.

Pridemore, W. A. (2014). The mortality penalty of incarceration: Evidence from a population-based case-control study of working-age males. *Journal of Health and Social Behavior 55*(2), 215–233.

Quetelet, A. (1984). *Adolphe Quetelet's research on the propensity for crime at different ages*. Anderson Publishing Company New York.

Rajkumar, A. S. and M. T. French (1997). Drug abuse, crime costs, and the economic benefits of treatment. *Journal of Quantitative Criminology 13*(3), 291–323.

Raphael, S. (2008). *Early Incarceration Spells and the Transition to Adulthood*, pp. 278–306. Russell Sage Foundation.

Raphael, S., H. J. Holzer, and M. A. Stoll (2006). Perceived criminality, criminal background checks, and the racial hiring practices of employers. *The Journal of Law and Economics 49*(2), 451–480.

Reynolds, A. J., J. A. Temple, S.-R. Ou, D. L. Robertson, J. P. Mersky, J. W. Topitzes, and M. D. Niles (2007). Effects of a school-based, early childhood intervention on adult health and well-being: A 19-year follow-up of low-income families. *Archives of Pediatrics & Adolescent Medicine 161*(8), 730–739.

Rosenbaum, P. R. and D. B. Rubin (1983, 04). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Rowe, D. C. and D. P. Farrington (1997). The familial transmission of criminal convictions. *Criminology 35*(1), 177–202.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers 3*(2), 135–146.

Saffer, H. and D. Dave (2002, January). Mental illness and the demand for alcohol, cocaine and cigarettes. Working Paper.

Schnittker, J. and A. John (2007). Enduring stigma: The long-term effects of incarceration on health. *Journal of Health and Social Behavior 48*(2), 115–130.

Schnittker, J., M. Massoglia, and C. Uggen (2012). Out and down: Incarceration and psychiatric disorders. *Journal of Health and Social Behavior 53*(4), 448–464.

Schubert, C. A., E. P. Mulvey, and C. Glasheen (2011). Influence of mental health and substance use problems and criminogenic risk on outcomes in serious juvenile offenders. *Journal of the American Academy of Child and Adolescent Psychiatry 50*(9), 925–937.

Seigle, E., N. Walsh, and J. Weber (2014). Core principles for reducing recidivism and improving other outcomes for youth in the juvenile justice system. Technical report, Council of State Governments Justice Center.

Skeem, J. L., C. Schubert, C. Odgers, E. P. Mulvey, W. Gardner, and C. Lidz (2006). Psychiatric symptoms and community violence among high-risk patients: A test of the relationship at the weekly level. *Journal of Consulting and Clinical Psychology 74*(5), 967–979.

Stürmer, T., R. Wyss, R. J. Glynn, and M. A. Brookhart (2014). Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. *Journal of Internal Medicine 275*(6), 570–580.

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

Tella, R. D. and E. Schargrodsky (2013). Criminal recidivism after prison and electronic monitoring. *Journal of Political Economy 121*(1), 28–73.

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

Todd, P. and W. Zhang (2019). A dynamic model of personality, schooling, and occupational choice. *Quantitative Economics*.

Uggen, C. and M. Thompson (2003). The socioeconomic determinants of ill-gotten gains: Within-person changes in drug use and illegal earnings. *American Journal of Sociology 109*(1), 146–185.

Umbach, R., A. Raine, and N. R. Leonard (2018). Cognitive decline as a result of incarceration and the effects of a cbt/mt intervention: A cluster-randomized controlled trial. *Criminal Justice and Behavior 45*(1), 31–55.

Wakefield, S. and C. Wildeman (2011). Mass imprisonment and racial disparities in childhood behavioral problems. *Criminology and Public Policy 10*(3), 793–817.

Wald, J. (2016, October). Arne's duncans proposal to redirect incarceration funds to education is right on the money.

Western, B. (2002). The impact of incarceration on wage mobility and inequality. *American sociological review 67*(4), 526–546.

Western, B. (2006). *Incarceration, Marriage, and Family Life*, pp. 131–167. Russell Sage Foundation.

Western, B., J. R. Kling, and D. F. Weiman (2001). The labor market consequences of incarceration. *Crime & Delinquency 47*(3), 410–427.

Wildeman, C. (2012). Imprisonment and (inequality in) population health. *Social Science Research 41*(1), 74–91.

Wilson, J. Q. and R. Herrnstein (1985). *Crime and human nature*. New York: Simon and Shuster.

Xie, Y., J. E. Brand, and B. Jann (2012). Estimating heterogeneous treatment effects with observational data. *Sociological Methodology 42*(1), 314–347. PMID: 23482633.

Xie, Y. and X. Wu (2005). Reply to jann: Market premium, social process, and statisticism. *American Sociological Review 70*(5), 865–870.

# Appendix A

# Chapter 2 Appendices

## A.1 Tables

Table A.1: Descriptive statistics-mean and standard deviation by sample.

| Variable | Full Sample | | Subsample for Estimation | | p-value |
|---|---|---|---|---|---|
| Race | | | | | |
| White | 0.202 | (0.402) | 0.257 | (0.438) | 0.007 |
| Black | 0.414 | (0.493) | 0.393 | (0.489) | 0.387 |
| Hispanic | 0.335 | (0.472) | 0.287 | (0.453) | 0.044 |
| Other | 0.048 | (0.214) | 0.063 | (0.243) | 0.175 |
| Philadelphia | 0.517 | (0.5) | 0.512 | (0.501) | 0.830 |
| Male | 0.864 | (0.343) | 0.812 | (0.391) | 0.003 |
| High School Diploma | | | | | |
| Mother | 0.515 | (0.5) | 0.581 | (0.494) | 0.009 |
| Father | 0.409 | (0.492) | 0.469 | (0.5) | 0.017 |
| Criminal Background | | | | | |
| Mother | 0.179 | (0.383) | 0.149 | (0.356) | 0.119 |
| Father | 0.351 | (0.477) | 0.307 | (0.462) | 0.069 |
| Complete Family | 0.147 | (0.354) | 0.185 | (0.389) | 0.035 |
| Cognitive Ability | 84.525 | (13.03) | 87.079 | (13.463) | 0.000 |
| Prior Detention | 0.493 | (0.5) | … | … | … |
| Prior Criminal Experience (years) | 1.205 | (1.492) | 0.564 | (1.059) | 0.000 |
| Initial Referral | | | | | |
| Person | 0.404 | (0.491) | 0.419 | (0.494) | 0.542 |
| Property | 0.252 | (0.434) | 0.234 | (0.424) | 0.425 |
| Drug | 0.155 | (0.362) | 0.155 | (0.363) | 0.999 |
| Other | 0.189 | (0.392) | 0.191 | (0.394) | 0.906 |
| Initial Disposition | | | | | |
| Residential Placement | 0.518 | (0.5) | 0.353 | (0.479) | 0.000 |
| Probation | 0.423 | (0.494) | 0.604 | (0.49) | 0.000 |
| Other | 0.043 | (0.203) | 0.030 | (0.17) | 0.200 |
| Baseline Interview | | | | | |
| Depression (std) | 0.251 | (1.162) | 0.257 | (1.232) | 0.925 |
| Anxiety (std) | 0.202 | (1.18) | 0.342 | (1.354) | 0.019 |
| Somatization (std) | 0.186 | (1.159) | 0.384 | (1.363) | 0.001 |
| Hostility (std) | 0.270 | (1.138) | 0.304 | (1.159) | 0.557 |
| Risk Perception | 52.794 | (29.028) | 54.954 | (27.823) | 0.141 |
| Temperance (std) | -0.254 | (0.994) | -0.263 | (0.958) | 0.852 |
| Substance Abuse | 0.621 | (0.485) | 0.604 | (0.49) | 0.485 |
| People | 1354 | | 303 | | |

Standard deviation in parentheses. P-value is calculated for the alternative hypothesis.

Table A.2: Propensity Score.

| Variable | Treatment | |
|---|---|---|
| Philadelphia | -0.320 | -0.184 |
| | (0.354) | (0.537) |
| Male | 1.327*** | 0.798 |
| | (0.415) | (1.620) |
| Race | | |
| White | -0.111 | -0.0603 |
| | (0.634) | (0.653) |
| Black | 0.388 | 0.226 |
| | (0.614) | (0.782) |
| Hispanic | -0.440 | -0.251 |
| | (0.619) | (0.838) |
| Cognitive Ability | 0.0180 | 0.0107 |
| | (0.151) | (0.152) |
| High School Diploma | | |
| Mother | -0.354 | -0.209 |
| | (0.301) | (0.527) |
| Mother: Missing Information | -0.755 | -0.445 |
| | (0.641) | (1.123) |
| Father | -0.516 | -0.300 |
| | (0.322) | (0.719) |
| Father: Missing Information | 0.0350 | 0.0194 |
| | (0.380) | (0.384) |
| Criminal Background | | |
| Mother | -0.0837 | -0.0557 |
| | (0.391) | (0.399) |
| Father | 0.183 | 0.110 |
| | (0.299) | (0.369) |
| Complete Family | -0.599* | -0.347 |
| | (0.363) | (0.832) |
| Initial Referral | | |
| Person | -0.348 | -0.191 |
| | (0.365) | (0.596) |
| Property | -0.292 | -0.157 |
| | (0.417) | (0.581) |
| Drug | -0.405 | -0.235 |
| | (0.476) | (0.695) |
| Prior Criminal Experience (years) | 0.183 | 0.105 |
| | (0.132) | (0.267) |
| Before the Treatment | | |
| Risk Perception | -0.00798 | -0.00466 |
| | (0.00535) | (0.0112) |
| Rewards of Crime | 0.0487 | 0.0282 |
| | (0.156) | (0.168) |
| Temperance | 0.0135 | 0.00949 |
| | (0.173) | (0.173) |
| Substance Abuse | 0.0501 | 0.0190 |
| | (0.318) | (0.331) |
| Depression | 0.139 | 0.0847 |
| | (0.166) | (0.232) |
| Anxiety | -0.150 | -0.0910 |
| | (0.184) | (0.253) |
| Somatization | 0.431** | 0.256 |
| | (0.170) | (0.543) |
| Hostility | -0.0451 | -0.0266 |
| | (0.170) | (0.178) |
| Propensity Score | … | 1.965 |
| | … | (5.866) |
| Constant | -0.978 | -1.596 |
| | (0.962) | (2.086) |
| Observations | 303 | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.3: Propensity score matching estimates of ATT of incarceration: Full subsample

| Matching algorithm | GSI | Depression | Anxiety | Somatization | Hostility |
|---|---|---|---|---|---|
| Propensity Score Matching | | | | | |
| 5 Nearest-neighbors | -0.056 | 0.1080 | -0.0720 | -0.1650 | 0.121* |
| | (0.151) | (0.106) | (0.132) | (0.162) | (0.071) |
| IPW | -0.027 | 0.0850 | -0.0890 | -0.2040 | 0.1410 |
| | (0.125) | (0.116) | (0.123) | (0.158) | (0.102) |
| IPW with RA | -0.066 | 0.0180 | -0.1360 | -0.2340 | 0.1350 |
| | (0.136) | (0.128) | (0.133) | (0.144) | (0.104) |
| Observations | | | 303 | | |
| Control | | | 161 | | |
| Treatment | | | 142 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.4: MDID estimates of ATT of incarceration: Full subsample

| Matching algorithm | GSI | Depression | Anxiety | Somatization | Hostility |
|---|---|---|---|---|---|
| Propensity Score Matching | | | | | |
| 5 Nearest-neighbors | -0.033 | 0.0160 | -0.0850 | -0.0890 | 0.2590 |
| | (0.130) | (0.212) | (0.248) | (0.261) | (0.216) |
| IPW | -0.051 | -0.0530 | -0.1680 | -0.278* | 0.1270 |
| | (0.144) | (0.132) | (0.138) | (0.144) | (0.12) |
| IPW with RA | … | … | … | … | … |
| | … | … | … | … | … |
| Observations | | | 303 | | |
| Control | | | 161 | | |
| Treatment | | | 142 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.5: Descriptive statistics-mean and standard deviation from control and treatment groups (by tercile).

| | Terciles | | | | | |
|---|---|---|---|---|---|---|
| | 1st | | 2nd | | 3rd | |
| Variables | Control | Treatment | Control | Treatment | Control | Treatment |
| Race | | | | | | |
| White | 0.329 | 0.300 | 0.296 | 0.278 | 0.172 | 0.176 |
| Black | 0.293 | 0.350 | 0.321 | 0.333 | 0.552 | 0.527 |
| Hispanic | 0.293 | 0.300 | 0.333 | 0.370 | 0.241 | 0.203 |
| Philadelphia | 0.427 | 0.550 | 0.469 | 0.463 | 0.621 | 0.595 |
| Male | 0.524 | 0.700 | 0.864 | 0.889 | 0.966 | 0.973 |
| High School Diploma | | | | | | |
| Mother | 0.561 | 0.650 | 0.654 | 0.519 | 0.448 | 0.568 |
| Father | 0.573 | 0.550 | 0.580 | 0.574 | 0.276 | 0.257 |
| Criminal Background | | | | | | |
| Father | 0.268 | 0.250 | 0.272 | 0.315 | 0.379 | 0.365 |
| Complete Family | 0.280 | 0.200 | 0.259 | 0.278 | 0.103 | 0.027 |
| Cognitive Ability | 86.805 | 90.650 | 88.247 | 86.519 | 85.655 | 86.041 |
| Prior Detention | … | … | … | … | … | … |
| Prior Criminal Experience (years) | 0.263 | 0.276 | 0.333 | 0.465 | 1.007 | 0.963 |
| Initial Referral | | | | | | |
| Person | 0.427 | 0.45 | 0.469 | 0.537 | 0.414 | 0.324 |
| Property | 0.305 | 0.3 | 0.259 | 0.241 | 0.207 | 0.189 |
| Drug | 0.146 | 0.15 | 0.136 | 0.093 | 0.103 | 0.176 |
| Initial Disposition | | | | | | |
| Residential Placement | 0.024 | 0.5 | 0.148 | 0.574 | 0.207 | 0.676 |
| Probation | 0.927 | 0.5 | 0.802 | 0.37 | 0.793 | 0.284 |
| Baseline Interview | | | | | | |
| Depression | -0.016 | -0.262 | 0.272 | 0.246 | 0.146 | 0.73 |
| Anxiety | 0.098 | -0.053 | 0.225 | 0.235 | 0.323 | 0.735 |
| Somatization | 0.009 | -0.234 | 0.185 | 0.248 | 0.308 | 0.872 |
| Hostility | 0.01 | 0.115 | 0.147 | 0.219 | 0.47 | 0.491 |
| Risk Perception | 67.613 | 68.07 | 56.316 | 52.487 | 40.445 | 43.319 |
| Temperance | -0.039 | -0.229 | -0.18 | -0.144 | -0.435 | -0.449 |
| Substance Abuse | 0.305 | 0.4 | 0.519 | 0.444 | 0.724 | 0.824 |
| People | 101 | | 101 | | 101 | |

Table A.6: Criminal engagement after the treatment.

| Variable | Probit | | Probit with IPW | |
|---|---|---|---|---|
| Philadelphia | -0.187 | -0.199 | -0.140 | -0.151 |
| | (0.223) | (0.224) | (0.374) | (0.376) |
| Male | 1.054*** | 1.051*** | 1.028** | 1.036** |
| | (0.232) | (0.232) | (0.453) | (0.453) |
| Race | | | | |
| White | -0.505 | -0.496 | -0.221 | -0.201 |
| | (0.401) | (0.401) | (0.700) | (0.700) |
| Black | 0.403 | 0.414 | 0.511 | 0.504 |
| | (0.398) | (0.397) | (0.683) | (0.682) |
| Hispanic | -0.406 | -0.396 | -0.116 | -0.135 |
| | (0.383) | (0.382) | (0.668) | (0.671) |
| Cognitive Ability | 0.252** | 0.248** | 0.249 | 0.244 |
| | (0.103) | (0.104) | (0.172) | (0.172) |
| High School Diploma | | | | |
| Mother | -0.632*** | -0.643*** | -0.593 | -0.595 |
| | (0.232) | (0.233) | (0.381) | (0.381) |
| Mother: Missing Information | -0.539 | -0.534 | -0.338 | -0.336 |
| | (0.369) | (0.368) | (0.665) | (0.666) |
| Father | -0.0434 | -0.0541 | 0.115 | 0.113 |
| | (0.231) | (0.232) | (0.376) | (0.377) |
| Father: Missing Information | -0.105 | -0.104 | -0.00237 | 0.0159 |
| | (0.280) | (0.281) | (0.448) | (0.451) |
| Criminal Background | | | | |
| Mother | 0.467* | 0.479* | 0.478 | 0.472 |
| | (0.264) | (0.265) | (0.461) | (0.463) |
| Father | 0.561*** | 0.556*** | 0.683* | 0.665* |
| | (0.213) | (0.214) | (0.353) | (0.356) |
| Complete Family | -0.757*** | -0.765*** | -0.621 | -0.627 |
| | (0.252) | (0.253) | (0.422) | (0.422) |
| Initial Referral | | | | |
| Person | 0.00583 | -0.0146 | 0.192 | 0.169 |
| | (0.269) | (0.271) | (0.433) | (0.438) |
| Property | 0.727** | 0.692** | 1.009** | 0.977* |
| | (0.306) | (0.311) | (0.505) | (0.514) |
| Drug | 0.507 | 0.468 | 0.728 | 0.688 |
| | (0.320) | (0.326) | (0.510) | (0.526) |
| 18 Years Old | 0.386** | 0.386** | 0.253 | 0.245 |
| | (0.193) | (0.194) | (0.317) | (0.318) |
| Prior Criminal Experience (years) | -0.129 | -0.124 | -0.160 | -0.163 |
| | (0.0938) | (0.0946) | (0.134) | (0.136) |
| Personal Characteristics (Previous Period) | | | | |
| Risk Perception | 0.00340 | 0.00324 | 0.000299 | -3.32e-05 |
| | (0.00373) | (0.00374) | (0.00619) | (0.00626) |
| Rewards of Crime | 0.191* | 0.188* | 0.215 | 0.213 |
| | (0.101) | (0.101) | (0.158) | (0.158) |
| Temperance | -0.351*** | -0.349*** | -0.313 | -0.310 |
| | (0.112) | (0.112) | (0.193) | (0.193) |
| Substance Abuse | -0.176 | -0.188 | -0.295 | -0.285 |
| | (0.206) | (0.209) | (0.328) | (0.334) |
| Mental Health (Previous Period) | | | | |
| Depression | 0.358*** | 0.348** | 0.458* | 0.452* |
| | (0.135) | (0.136) | (0.234) | (0.234) |
| Anxiety | -0.257 | -0.257 | -0.298 | -0.293 |
| | (0.161) | (0.161) | (0.273) | (0.273) |
| Somatization | 0.163 | 0.157 | 0.189 | 0.202 |
| | (0.137) | (0.140) | (0.232) | (0.238) |
| Hostility | 0.329*** | 0.331*** | 0.301 | 0.288 |
| | (0.123) | (0.124) | (0.226) | (0.229) |
| Year Dummies | | | | |
| 2002 | -0.309 | -0.317 | -0.264 | -0.275 |
| | (0.252) | (0.252) | (0.371) | (0.372) |
| 2003 | -0.198 | -0.192 | -0.0885 | -0.0945 |
| | (0.210) | (0.210) | (0.355) | (0.358) |
| Education | -0.00420 | -0.00533 | -0.0224 | -0.0230 |
| | (0.0118) | (0.0119) | (0.0207) | (0.0208) |
| Experience (Legal Sector) | -0.00881 | -0.0101 | -0.0132 | -0.0137 |
| | (0.0176) | (0.0177) | (0.0307) | (0.0307) |
| Monetary Benefits (Previous Period) | | -0.000264 | | -0.000338 |
| | | (0.000626) | | (0.000927) |
| Monetary Benefits: Missing | | -0.337 | | -0.146 |
| | | (0.454) | | (0.749) |
| Constant | -0.801 | -0.412 | -0.970 | -0.759 |
| | (0.648) | (0.830) | (1.114) | (1.410) |
| Observations | | 321 | | |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

## A.2 Figures

Figure A.1: Trent before the treatment: Depression.

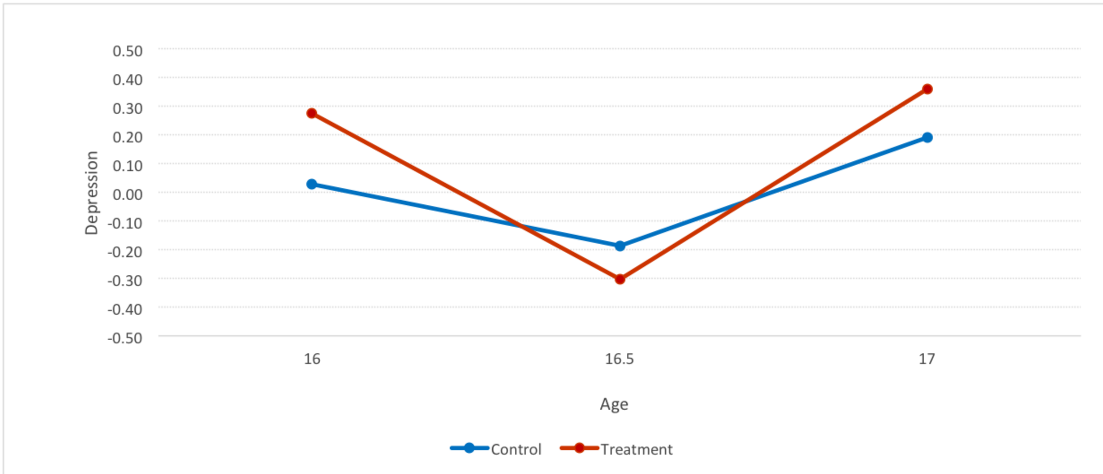



Figure A.2: Trent before the treatment: Anxiety.

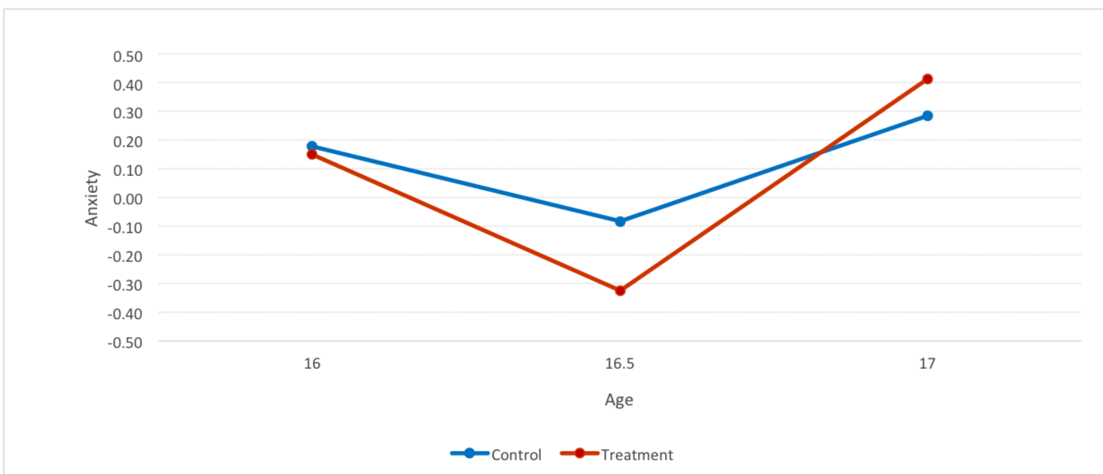


Figure A.3: Trent before the treatment: Somatization.

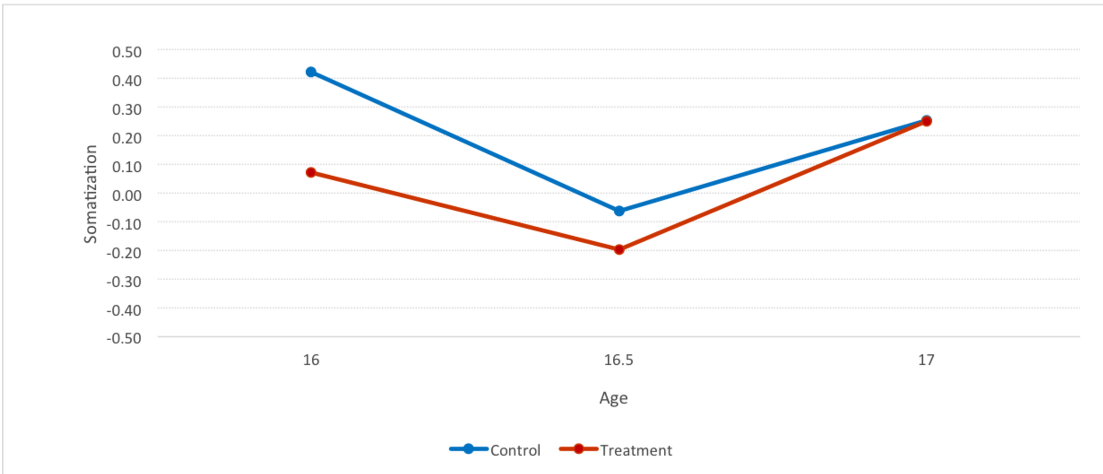
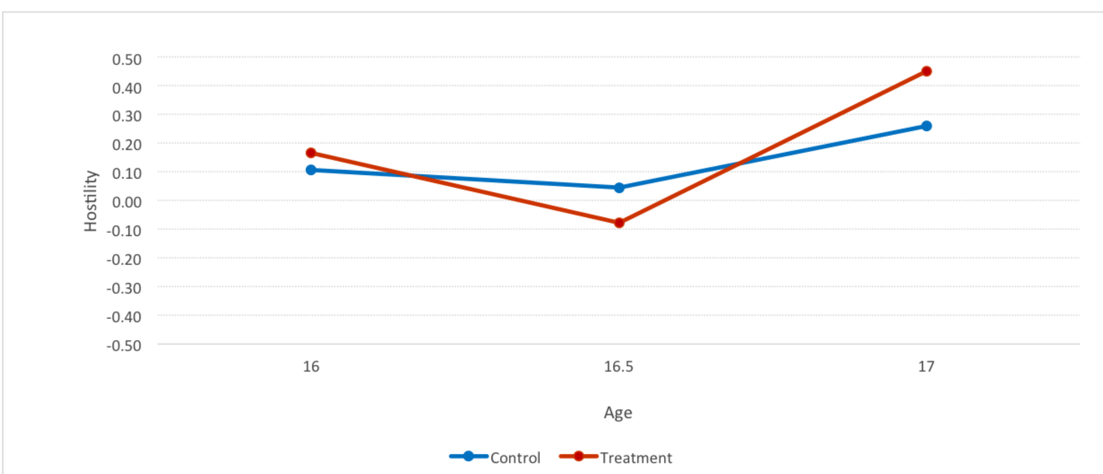



Figure A.4: Trent before the treatment: Hostility.



# Appendix B

# Chapter 3 Appendices

## B.1 Tables

Table B.1: Estimated parameters from probit models for crime, high school graduation and home production.

| Variable | Income crime | All crime | Home production | High school graduation |
|---|---|---|---|---|
| Mental health (previous period) | | | | |
| Depression | -0.012 | -0.004 | -0.012 | -0.039*** |
| | (0.012) | (0.015) | (0.010) | (0.011) |
| Anxiety | 0.005 | -0.015 | 0.004 | 0.011 |
| | (0.014) | (0.016) | (0.011) | (0.012) |
| Somatization | -0.002 | -0.004 | 0.007 | -0.007 |
| | (0.012) | (0.014) | (0.009) | (0.011) |
| Hostility | 0.048*** | 0.084*** | -0.000 | 0.025*** |
| | (0.011) | (0.014) | (0.009) | (0.009) |
| Self-control (previous period) | -0.068*** | -0.065*** | -0.001 | 0.033*** |
| | (0.010) | (0.012) | (0.008) | (0.008) |
| Observations | 2,859 | 2,859 | 2,859 | 2,431 |

Notes: Standard errors in parentheses. Each measure was standardized to have mean 0 and variance 1. Observed individual characteristics include: age, age squared, location, race, IQ, accumulated periods of education, legal experience and criminal experience. Periods of sector-specific experience and education are measured in semesters. Family background includes: mother with high school, father with high school, mother with incarceration records, father with incarceration records, and an indicator for being raised in a complete family. Criminal history icludes: type of crime for the initial referral (e.g. person, property, drug, other), proxy for unobserved criminal experience, and a dummy for having incarceration records so far.
*** p<0.01, ** p<0.05, * p<0.1

# Appendix C

# Chapter 4 Appendices

## C.1 Pathways to Desistance: Auxiliary Models

Table C.1: Linear probability models for choices in freedom.

| Variables | (1) Employment | (2) Crime | (3) School | (4) Unemployment |
|---|---|---|---|---|
| Age | -0.0170 | 0.00802 | 0.00436 | 0.00462 |
| | (0.0193) | (0.0213) | (0.0343) | (0.0241) |
| Experience | | | | |
| Legal | 0.0769*** | -0.00918*** | -0.0108*** | -0.0569*** |
| | (0.00382) | (0.00194) | (0.00244) | (0.00344) |
| Criminal | -0.0271*** | 0.0911*** | -0.00997*** | -0.0541*** |
| | (0.00649) | (0.00580) | (0.00338) | (0.00553) |
| Education | 0.0145*** | -0.00410** | 0.00896*** | -0.0193*** |
| | (0.00348) | (0.00192) | (0.00268) | (0.00285) |
| Trainning | 0.0234*** | -0.000150 | -0.0163*** | -0.00702* |
| | (0.00459) | (0.00237) | (0.00282) | (0.00418) |
| Mental health | -0.0218*** | 0.0162*** | -0.00406 | 0.00970* |
| | (0.00609) | (0.00443) | (0.00473) | (0.00554) |
| Self-control | 0.00485 | -0.0188*** | 0.00968** | 0.00423 |
| | (0.00572) | (0.00360) | (0.00444) | (0.00489) |
| Cognitive | 0.0250*** | 0.000177 | -0.00375 | -0.0214*** |
| | (0.00576) | (0.00340) | (0.00445) | (0.00505) |
| Lagged school | -0.134*** | -0.0318*** | 0.269*** | -0.103*** |
| | (0.0166) | (0.00910) | (0.0141) | (0.0133) |
| Age dummies | | | | |
| 16-18 | -2.704*** | -0.0497 | 3.514*** | -0.761 |
| | (0.642) | (0.669) | (1.067) | (0.755) |
| More than 18 | 0.730 | 0.480 | -0.493 | -0.717 |
| | (0.595) | (0.651) | (1.049) | (0.742) |
| Detention so far | -0.0309** | 0.0232*** | -0.0139 | 0.0217* |
| | (0.0143) | (0.00752) | (0.0112) | (0.0116) |
| Interaction terms | | | | |
| 16-18 | 0.0852*** | 0.00152 | -0.111*** | 0.0239 |
| | (0.0207) | (0.0218) | (0.0348) | (0.0246) |
| More than 18 | -0.00871 | -0.0132 | -0.00232 | 0.0242 |
| | (0.0194) | (0.0213) | (0.0344) | (0.0242) |
| Constant | 0.676 | -0.195 | 0.508 | 0.0114 |
| | (0.588) | (0.646) | (1.045) | (0.736) |
| Observations | 5,888 | 5,888 | 5,888 | 5,888 |
| R-squared | 0.241 | 0.178 | 0.400 | 0.100 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.2: Linear probability models for choices in detention.

| Variables | (1) Crime | (2) Trainning |
|---|---|---|
| Age | -0.00657 | 0.00657 |
| | (0.0305) | (0.0305) |
| Experience | | |
| Legal | -0.00521 | 0.00521 |
| | (0.00452) | (0.00452) |
| Criminal | 0.0323*** | -0.0323*** |
| | (0.00569) | (0.00569) |
| Education | -0.000674 | 0.000674 |
| | (0.00367) | (0.00367) |
| Trainning | 0.00128 | -0.00128 |
| | (0.00299) | (0.00299) |
| Penalty length | -0.000792 | 0.000792 |
| | (0.00223) | (0.00223) |
| Mental health | 0.0246*** | -0.0246*** |
| | (0.00678) | (0.00678) |
| Self-control | -0.0124** | 0.0124** |
| | (0.00550) | (0.00550) |
| Cognitive | 0.00969* | -0.00969* |
| | (0.00554) | (0.00554) |
| Lagged school | -0.134*** | 0.134*** |
| | (0.0183) | (0.0183) |
| Juvenile | -0.0337* | 0.0337* |
| | (0.0176) | (0.0176) |
| Age dummies | | |
| 16-18 | -0.356 | 0.356 |
| | (0.943) | (0.943) |
| More than 18 | -0.144 | 0.144 |
| | (0.938) | (0.938) |
| Interaction terms | | |
| 16-18 * age | 0.0104 | -0.0104 |
| | (0.0306) | (0.0306) |
| More than 18 * age | 0.00426 | -0.00426 |
| | (0.0305) | (0.0305) |
| Constant | 0.398 | 0.602 |
| | (0.937) | (0.937) |
| Observations | 2,601 | 2,601 |
| R-squared | 0.119 | 0.119 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.3: Mincer regression: Legal sector.

| Variables | (1) Log wages |
|---|---|
| Education | 0.0183*** |
| | (0.00378) |
| Trainning | 0.0125** |
| | (0.00530) |
| Experience | 0.0932*** |
| | (0.0138) |
| Experience Squared | -0.00360** |
| | (0.00158) |
| Mental health | 0.00651 |
| | (0.0109) |
| Self-control | 0.0122 |
| | (0.00845) |
| Cognitive | 0.00236 |
| | (0.00950) |
| Detention so far | 0.0418** |
| | (0.0187) |
| Lagged employment | 0.0145 |
| | (0.0222) |
| Constant | 7.049*** |
| | (0.0305) |
| Observations | 2,842 |
| R-squared | 0.131 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.4: Mincer regression: Criminal sector.

| Variables | (1) Log illegal earnings |
|---|---|
| Criminal experience | 0.209** |
| | (0.0830) |
| Criminal experience squared | -0.0203* |
| | (0.0108) |
| Mental health | -0.0654 |
| | (0.0413) |
| Self-control | -0.000473 |
| | (0.0589) |
| Cognitive | -0.0957* |
| | (0.0513) |
| Detention so far | 0.804*** |
| | (0.143) |
| Lagged crime | 0.0480 |
| | (0.130) |
| Constant | 6.893*** |
| | (0.172) |
| Observations | 483 |
| R-squared | 0.113 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.5: Personal capabilities: Law of motion.

| Variables | (1) Mental health | (2) Self-control |
|---|---|---|
| Autoregressive term | 0.895*** | 0.925*** |
| | (0.00463) | (0.00417) |
| Dummies for detention status | | |
| Juvenile facility | 0.113*** | -0.0483*** |
| | (0.0163) | (0.0146) |
| Adult facility | 0.0480*** | -0.0396 |
| | (0.0112) | (0.0101) |
| Constant | -0.00461 | 0.0193*** |
| | (0.00557) | (0.00501) |
| Observations | 8,489 | 8,489 |
| R-squared | 0.816 | 0.855 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

## C.2 Additional Simulations

Figure C.1: Evolution of human and criminal capital for marginal permanent offenders.

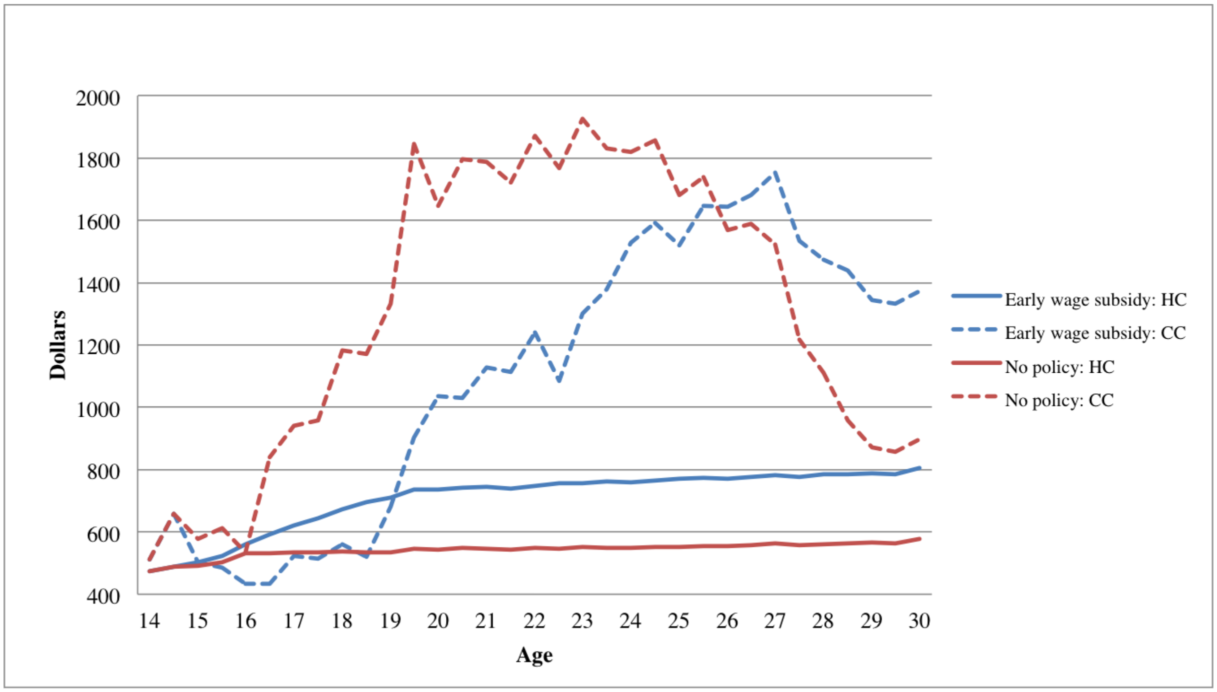


Figure C.2: Evolution of human and criminal capital for marginal permanent offenders.

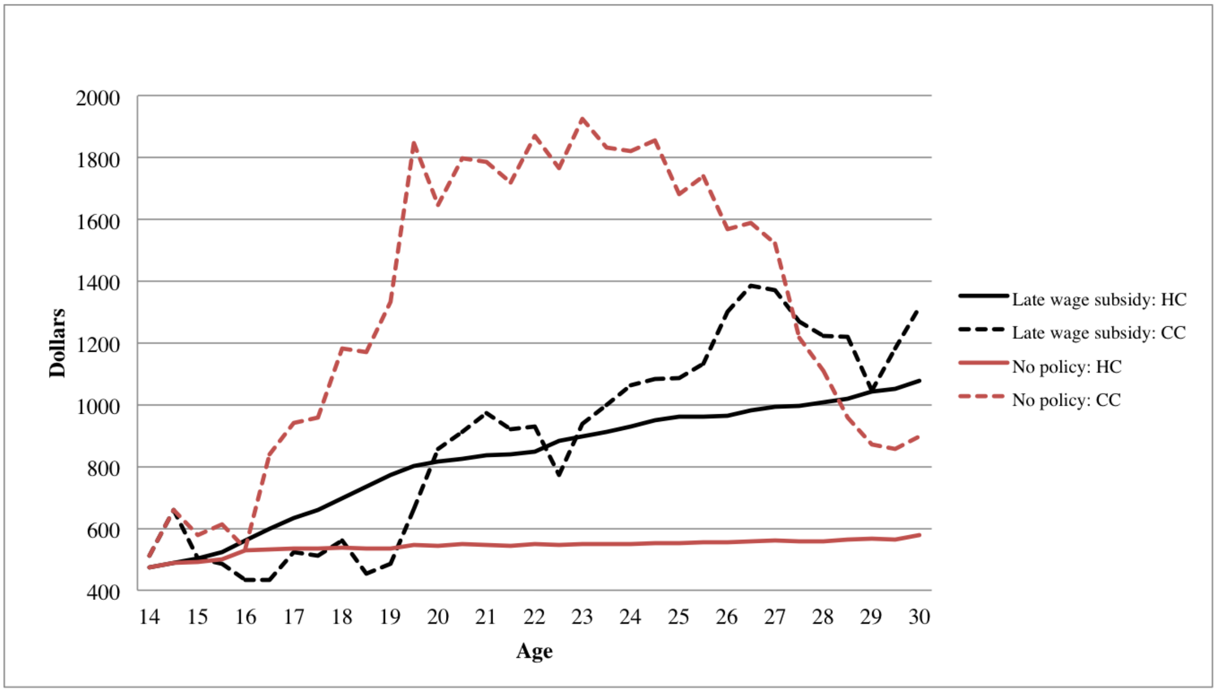



Figure C.3: Evolution of human and criminal capital for permanent offenders.

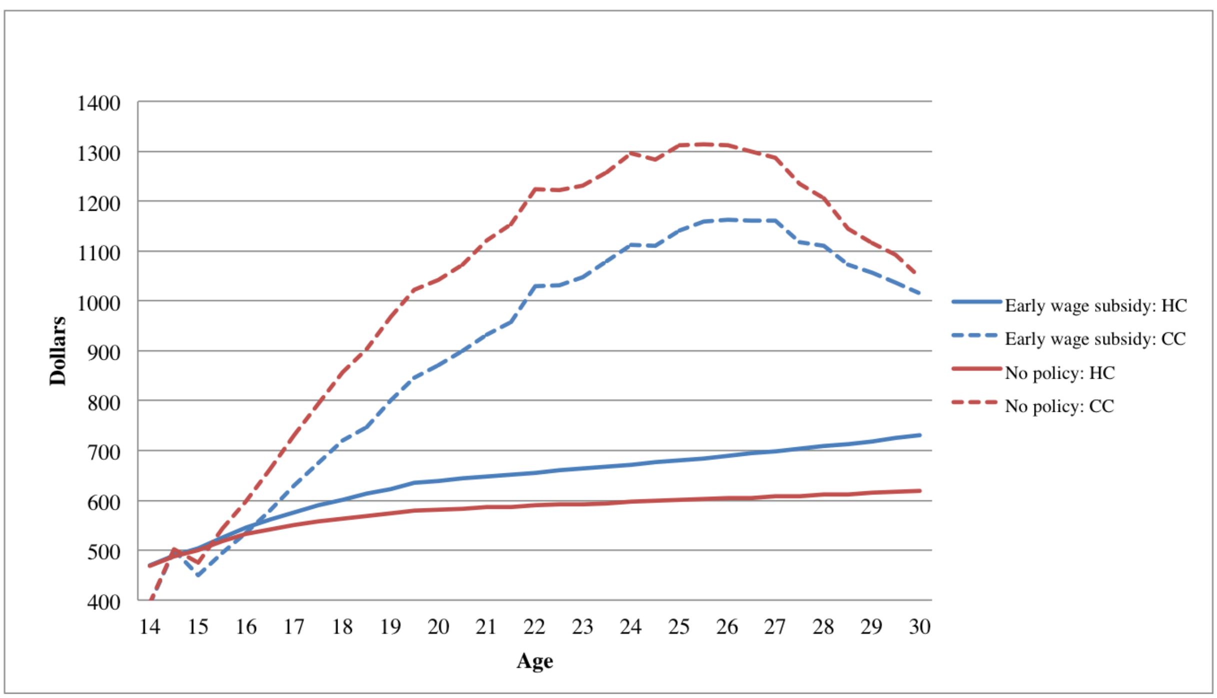


Figure C.4: Evolution of human and criminal capital for permanent offenders.

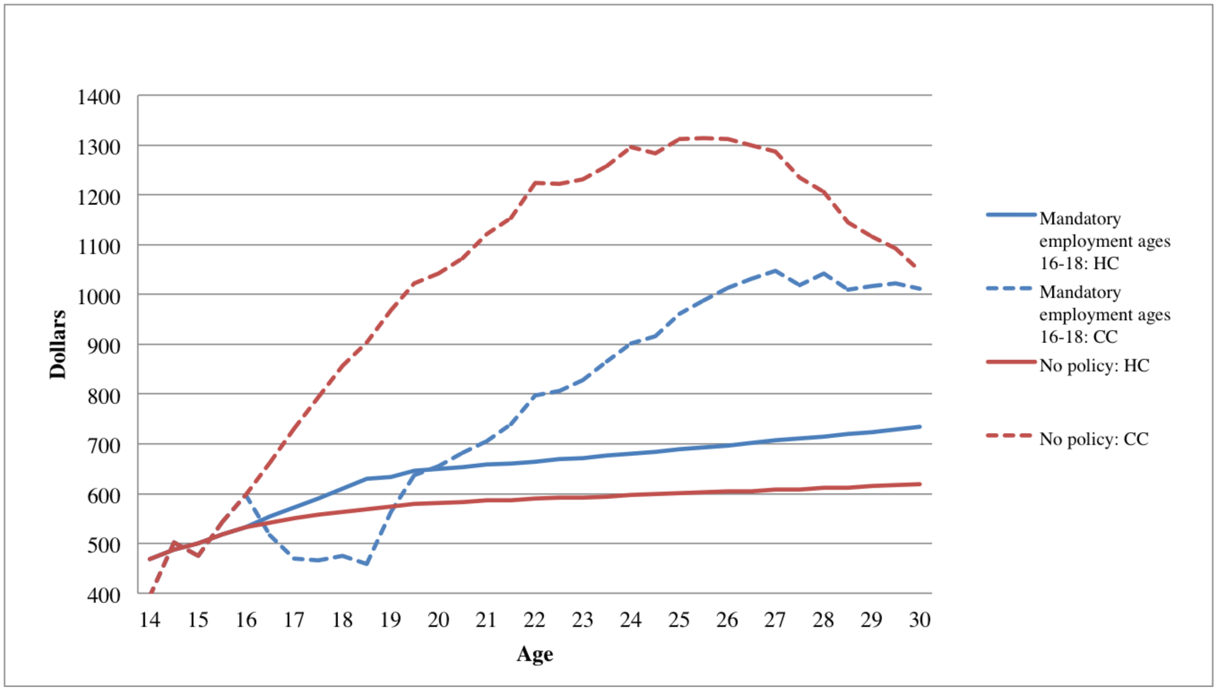



Figure C.5: Evolution of human and criminal capital for permanent offenders.

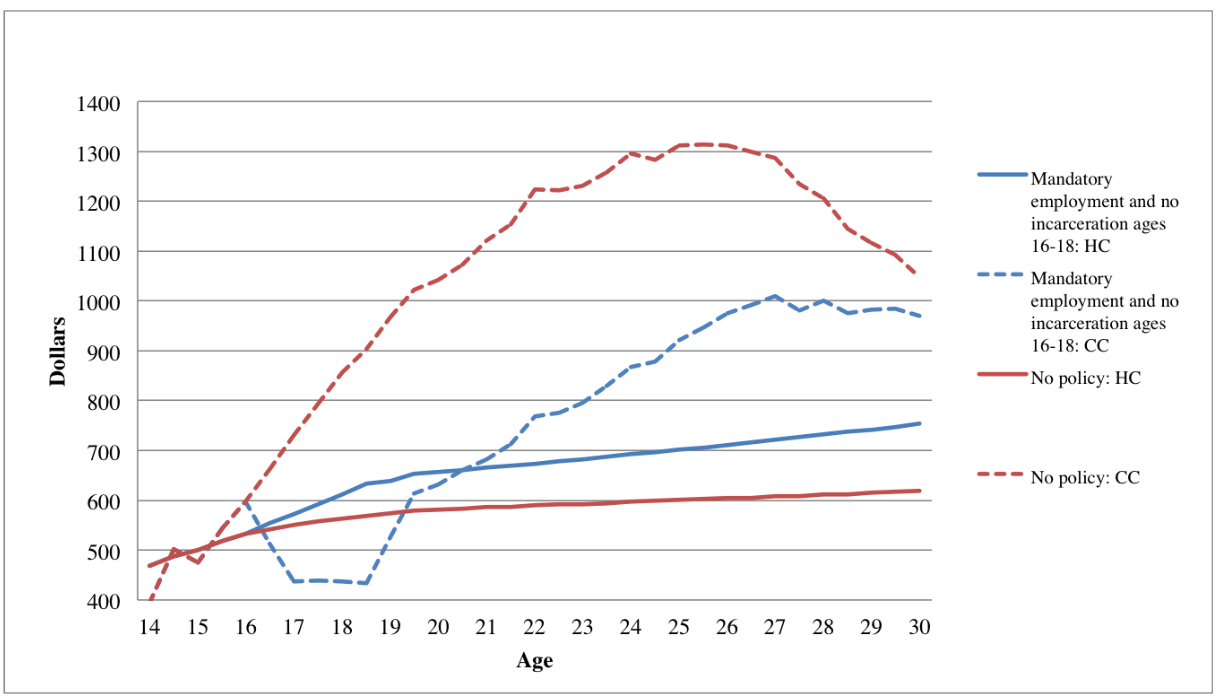


Figure C.6: Evolution of human and criminal capital for permanent offenders.

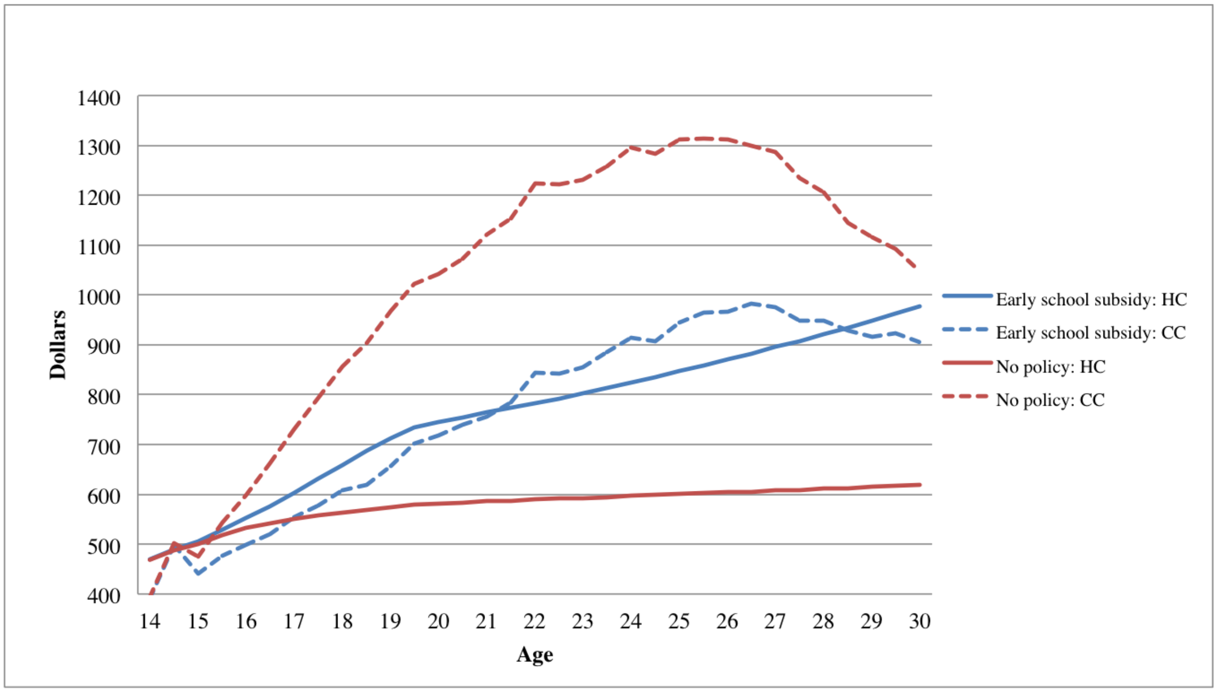



Figure C.7: Evolution of human and criminal capital for permanent offenders.

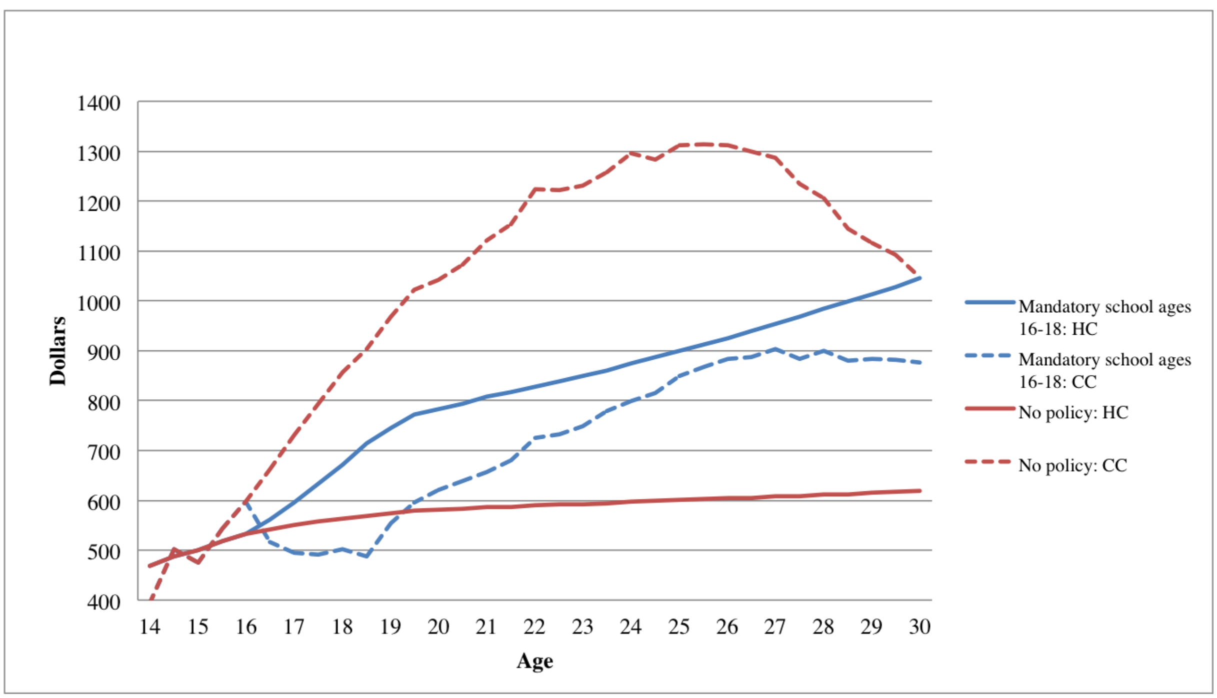
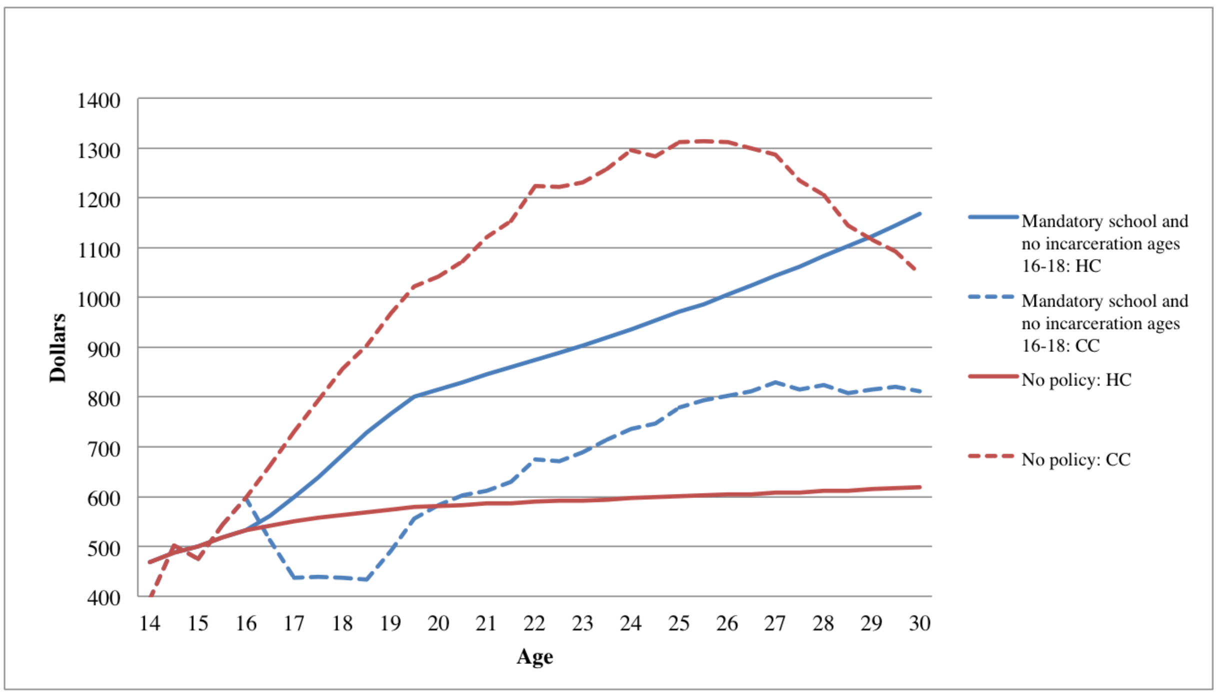


Figure C.8: Evolution of human and criminal capital for permanent offenders.



# Curriculum Vitae

| | |
|---|---|
| **Name:** | Diego F. Salazar |
| **Post-Secondary Education and Degrees:** | The University of Western Ontario<br>London, Ontario, Canada<br>2012-2019, Ph.D. Economics |
| | Universidad de los Andes<br>Bogota, Colombia<br>2005-2007, M.A. Economics |
| | Universidad de los Andes<br>Bogota, Colombia<br>2000-2004, B.S. Economics |
| **Honours and Awards:** | Western Graduate Research Scholarship<br>The University of Western Ontario<br>2012-2016 |
| | Graduate Tutorial Leader of the Year<br>The University of Western Ontario<br>2013 |
| **Related Work Experience:** | Course Instructor<br>The University of Western Ontario<br>2014-2018 |
| | Teaching Assistant<br>The University of Western Ontario<br>2012-2016 |
| | Research Assistant<br>The University of Western Ontario<br>2014-2017 |