DISENTANGLING CRIMINAL CAREERS
FOR DISADVANTAGED YOUTHS
This thesis was sponsored by the Netherlands Institute for the Study of Crime and Law Enforcement and the Phoolan Devi Institute of the VU University Amsterdam.
DISSENGAING CRIMINAL CAREERS
FOR DISADVANTAGED YOUTHS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Rechtsgeleerdheid
op vrijdag 16 januari 2015 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door
Geert Mesters
geboren te Utrecht
promotoren:  prof.dr. mr. C.C.J.H. Bijleveld
            prof.dr. S.J. Koopman

copromotor:  dr. V.R. van der Geest
Acknowledgments

The writing of this thesis has been a very interesting and exciting project. From my first day until my last day the NSCR has provided a great environment for doing research. The excellent support staff and the availability of unique data has ultimately made it possible for me to write this thesis.

Many people from inside and outside the NSCR have contributed to this thesis. First, I mention my supervisors Catrien Bijleveld and Siem Jan Koopman. Both were always encouraging and generous with their time. When I first came up with the idea to submit two dissertations they were immediately supportive and helped me in every possible way to achieve this goal. Next to my supervisors Victor van der Geest, in his role as co-supervisor, has helped a lot with writing the theoretical parts of this thesis.

Second, I feel lucky to have met a lot of great people at the NSCR. A special thanks is reserved for Rinke and Steve with whom I have shared great trips to Budapest and recently to Barcelona. I hope we will have many more of these fun weekends. Also I would like to thank the pubquiz team, the people from the department of criminology at the VU University Amsterdam and all other people from the NSCR.

Third, I would like to thank all my friends, parents and sister who have always supported me in a lot of different ways. Finally, I sincerely hope that this thesis can help to improve the lives of disadvantaged youths.

Barcelona, November 2014
## Contents

1 General introduction ........................................... 1
   1.1 Introduction .............................................. 2
   1.2 Conceptual framework ..................................... 5
   1.3 Theoretical foundation .................................... 7
      1.3.1 Criminological and economic theories ............. 8
      1.3.2 The role of skills and personality traits .......... 10
   1.4 Towards a testable empirical framework ................. 11
   1.5 Data summary ............................................. 14
      1.5.1 17up study ........................................ 14
      1.5.2 TRANS-5 study .................................... 17
   1.6 Summary of the remainder ................................ 18

2 Explaining the Age-Crime Curve of Disadvantaged Youths by Dynamic Payoffs from Cognitive and Social Skills 21
   2.1 Introduction .............................................. 21
   2.2 Explaining the age-crime curve .......................... 24
   2.3 The current study ......................................... 26
   2.4 Statistical model ......................................... 26
   2.5 Data description .......................................... 28
      2.5.1 Cognitive and social skill variables ............. 29
      2.5.2 Criminal records data ............................... 30
      2.5.3 Control variables ................................... 31
   2.6 Results ................................................... 32
      2.6.1 Payoffs from childhood skills ...................... 32
4.4.1 Age-invariant effects ................................................. 98
4.4.2 Age-varying effects .................................................. 103
4.5 Policy implications ..................................................... 105
4.6 Conclusion ............................................................... 111

5 The Effect of Unemployment on Crime in High Risk Families in the Netherlands between 1920 and 2005 ......................................................... 121
5.1 Introduction .............................................................. 121
5.2 Data ................................................................. 124
  5.2.1 Conviction data ..................................................... 125
  5.2.2 Additional family information ................................. 128
  5.2.3 Relation to the national unemployment rate ................. 129
5.3 Empirical panel data strategy ....................................... 131
  5.3.1 The observation model .......................................... 132
  5.3.2 The signal ....................................................... 133
  5.3.3 Results .......................................................... 135
5.4 Time-varying unemployment effects .................................. 141
  5.4.1 Results .......................................................... 141
5.5 Conclusion ............................................................... 144

6 General conclusion ....................................................... 151
6.1 Main empirical findings .............................................. 152
6.2 Implications for life-course theories ............................... 155
  6.2.1 An age-graded perspective for skills ......................... 158
6.3 Implications for public policy ....................................... 160
  6.3.1 Individual-level interventions ................................. 160
  6.3.2 More on government policies ................................. 164
6.4 Comparison to traditional modeling approaches .................. 165
6.5 Directions for future research ...................................... 167

Bibliography .............................................................. 169
  References ............................................................ 169

Samenvatting (Summary in Dutch) ...................................... 185
Chapter 1

General introduction

Suppose that you were born in a disadvantaged socioeconomic environment where your parents are unable or unwilling to invest sufficiently in you during childhood. In your family alcohol abuse, drug use and unemployment are common, making it questionable whether your early childhood alone was all that pleasant. As a result of your disrupted family life you are experiencing problems adjusting at school. Your grades are low and you are often sent to the principal’s office. Your personal problems are not confined to school, also at home things are not going great. You are constantly fighting with your parents and rather spend as little as possible time at home. Instead of fixing your problems it seems a better idea to hang out with your friends and have fun. Since you are under age most of the hanging out takes place on the street. There you often get into trouble and at some occasions the police shows up. Eventually, during a particular wild evening you go too far and get arrested. No big deal, most of your friends have been arrested before and the first time the judge will go easy on you. Unfortunately, this arrest is not the only one. Over the next few months you get arrested a few more times and eventually the magistrate of the juvenile court decides that you have to go to a juvenile treatment facility. In the facility you receive treatment for your behavioral problems and follow classes to increase your education level. In your late teenage years you have to leave the treatment facility and you are put on a train heading into adulthood.
1.1 Introduction

In this thesis we empirically investigate to what extent such a detrimental start to “adult” life, as sketched above, has lasting impacts on socioeconomic adult outcomes, and whether life course transitions, such as those from employment and intimate relationships, can remain to alter adult life outcomes. The goal is to provide a conceptual and empirical framework for studying the effects of childhood outcomes on a variety of adult outcomes, while acknowledging that transitions during adulthood may influence future adult outcomes. The main focus in this thesis is on explaining adulthood offending for disadvantaged youths, but a variety of other socioeconomic outcomes, such as employment, social welfare, drug use and intimate relationships, are additionally studied.

The phrasing of the purpose of this thesis may suggest that the possible consequences of a disadvantaged childhood are either black or white. Black being that childhood entirely determines socioeconomic adult outcomes. This perspective is reflected in the old Jesuit motto: “Give me a child until he is seven and I will give you the man”. The motto emphasizes the believe of Jesuit missionaries that when they were given a child at young age they could form the child into a “good” Christian for the remainder of its life. In contrast to the black perspective, the white perspective suggests that adulthood outcomes are determined independent from childhood circumstances. Clearly, both black and white perspectives are likely to have some merit and the goal is to assess their relative explanatory power (Nagin & Paternoster, 1993; Blokland & Nieuwbeerta, 2010).

While the separation of effects from childhood and adulthood is of theoretical interest, developing public policies to improve adult outcomes for disadvantaged youths requires additional insights. In particular, it is important to determine at which stage during life what kind of interventions in the form of investments in positive adult life outcomes are most effective and efficient. Many interventions may improve adult life outcomes and this thesis focuses on a subset of these. In particular, we are mainly interested in interventions that improve cognitive and social skills, and interventions that facilitate employment. The term “efficiency” for the interventions is to be seen in

---

1 The quote is attributed to St. Francis Xavier a Roman Catholic Missionary (7 April 1506 - 3 December 1552).
CHAPTER 1: GENERAL INTRODUCTION

a broad perspective where not only monetary aspects are important, but also societal impacts from crime reductions and drugs use are deemed relevant (Cohen, 1998).

The development of a framework that is able to explain socioeconomic adult outcomes in terms of childhood factors and influences from other adult outcomes is a challenging task. This thesis takes a step-by-step approach and relies on insights from criminology, sociology, economics and psychology. In particular, the conceptual framework is based on theories from the aforementioned research areas, which are translated into an empirical framework using mathematical language. The translation facilitates the testing of the framework which relies on observational data and a variety of econometric methods\(^2\).

In the first step we develop a model for incorporating childhood skills, which are arguably partially determined by personality traits, into a model for adulthood offending. The childhood skills include cognitive skills and a broad array of social skills. While the role of cognitive ability has traditionally been of major interest in social sciences, the role of social skills has received less attention in explaining offending outcomes (Hill, Roberts, Grogger, Guryan, & Sixkiller, 2011). The model incorporates the intuitively appealing notions that multiple skills are important in explaining adult offending and that different skills can be important at different stages in life (Cunha & Heckman, 2007).

Second, a framework is developed that acknowledges that adult outcomes interact with each other over the adult life span. In particular, adult outcomes for offending, employment, social welfare, and many others, may influence each other over the life course in both directions. For example, offending may reduce future employment probabilities and employment may reduce offending (Lageson & Uggen, 2013). Such bi-directional relationships need to be disentangled in order to be able to assess the relative returns from investments in adult life transitions. Some of the adult outcomes are made possible by the economic and political institutions of the Netherlands, the country in which the empirical study takes place. In particular, social welfare out-

\(^2\)We adopt the following definitions for the conceptual and empirical frameworks. The conceptual framework is the structure of assumptions, principles, and rules that holds together the decomposition of the adult outcomes into childhood and adulthood factors. The empirical framework, or model, is the translation of the conceptual framework into mathematical equations. The unknown structural model parameters of the mathematical equations, which govern the signs and magnitudes of the relationships between the endogenous and exogenous model components, can be directly estimated from the empirical framework when the model components are approximated by observational data.
comes are facilitated by the government of the Netherlands and a “choice” to rely on welfare benefits is thus facilitated by the government. In this sense the framework also provides an assessment for externalities that are created by social welfare policies.

Third, when quantifying the impact of childhood skills on adult outcomes at different stages of adulthood, the aforementioned adult life interactions need to be incorporated. Adult life transitions may act as multipliers that increase or decrease the payoffs from investments in childhood skills. For example, suppose that early interventions aimed at improving social skills reduce offending during adolescence, then the reduction in offending may increase subsequent employment probabilities, over and above the marginal influence of any social skills intervention on employment probabilities. The formulation of this developmental conceptual framework, its translation into an empirical framework and the subsequent testing of the empirical framework constitutes the main contribution of this thesis to the developmental perspective on crime.

The fourth study that is included in this thesis takes a historical perspective for the relationship between employment and crime. The main question that we study is to what extent the effect of unemployment on crime has changed during the last century. While this question is interesting by itself, within the context of the thesis the fourth study is to be regarded as a warning that historical context matters. In particular, it is questionable whether the same childhood skills and adult life transitions that were important in the past will remain to be important in the future. This implies that continuous updating of the relevant components of the framework is required for adequately adjusting public policies for disadvantaged youths.

We test the empirical framework using data for samples of disadvantaged youths who were institutionalized in a juvenile treatment facility in the Netherlands. We include two samples, males and females, that were released in the 1990’s and a sample of males that were released in the 1910’s. The samples compromise a segment of the Dutch population that is overrepresented in offending statistics (Boendermaker, 1999). We consistently refer to the youths as disadvantaged, while we acknowledge that other characterizations such as vulnerable youths or high risk youths are equally applicable.

In total around 4,000 youths are institutionalized in a criminal justice or youth care institution every year in the Netherlands (CBS, 2013). Based on their early encounters with the justice system and/or their behavioral problems, the youths can be regarded
as pertaining to a disadvantaged subgroup of youths who are at high risk for offending, have low employment participation rates and are relatively often the recipients of social welfare payments (van der Geest, 2011; Mesters, van der Geest, & Bijleveld, 2014; Verbruggen, 2014). While in an institution their behavioral problems are treated and low-level education is provided. In their late teenage years these youths typically leave the treatment facility and their “adult” life starts. Given their troubled backgrounds they are likely to experience problems adapting to adult life roles (Osgood, Foster, Flanagan, & Ruth, 2005). Our goal is to assess what kinds of interventions can best be made to improve the socioeconomic adult outcomes of such disadvantaged youths.

The remainder of this general introduction is organized as follows. In Section 1.2 we discuss a stylized version of the conceptual framework that is developed in this thesis. Section 1.3 links this framework to various criminological, sociological, psychological and economic theories, and Section 1.4 outlines the steps that facilitate the translation of the conceptual framework in an empirical framework. In Section 1.5 we discuss the data that we use to test the framework. The remainder of this thesis is summarized in Section 1.6.

1.2 Conceptual framework

In Figure 1.1 we show a stylized version of the conceptual framework that is developed in this thesis. The figure shows a three-stage developmental process where childhood is summarized by one period and adulthood compromises two periods. In our empirical studies adulthood is divided in many more periods, but the main concepts can be explained by a two-period framework for adulthood.

During childhood individuals develop skills, which in the subsequent chapters are split into cognitive and a variety of social skills. These skills are formed by a hierarchical dynamic process where parental investments and initial birth endowments are the inputs (Heckman, 2006; Cunha & Heckman, 2007). The childhood signals, which for example may include education levels and information from criminal records, are partially determined by the skills. In particular, cognitive skills may increase education levels and social skills may reduce the probability for a juvenile criminal record. The signals provide additional information based on which potential employers, potential romantic partners and law enforcement agencies make decisions (e.g., Weiss,
Since such entities typically do not have the resources to assess skills themselves, they may rely on signals instead when making decisions.

Given the skills and signals, the youths enter the first adulthood period. In this period all childhood skills and signals may influence the adult outcomes (Heckman, Stixrud, & Urzúa, 2006). We include adult outcomes for offending, employment, social welfare, drug use and intimate relationships. We note that in empirical applications there are always unobserved variables that may influence the adult outcomes. To correct for these influences we use statistical control variables which we discuss in the outline of the empirical framework in Section 1.4.

Next, the youths transition to the second adulthood period. The outcomes in this period are determined by (a) the persistent effects from childhood skills and signals, and (b) the effects from the adult outcomes from the first adulthood period. The effects of the skills and signals are allowed to persist in the second adult period but their payoff may be different. In particular, the arrow that directly relates the skills and signals to the second adult period may imply different signs and magnitudes for the effects of the skills and signals. This construction acknowledges that many skills are formed early in life and may often be considered stable in their ranking thereafter, but their mean-level impact may vary with age (Roberts & DelVecchio, 2000; Cunha & Heckman, 2007; Blonigen, 2010). For example, cognitive ability is found to be stable from age 10 onwards in Hopkins and Brecht (1975), but the impact of cognitive ability on adult outcomes can be different at ages 18 and 30.

The effects from the first adulthood period on the second adulthood period are typically referred to as structural effects in the sense that the first period outcomes have predictive ability for the second period outcomes over and above the effects of the skills and signals. For example, employment may structurally lower future offending probabilities as it increases the costs associated with offending (Lochner, 2004). In general, all adult outcomes from period one may have structural effects on all adult outcomes in period two. The testing of the empirical framework aims to assess the signs and magnitudes of the structural effects.

---

3The distinction between structural and spurious effects stems from Heckman (1981a) and Heckman (1981c). Spurious effects are in this sense defined as effects stemming from variables, other than the adult outcome variables, that create correlation between the adult outcome variables in periods one and two. These include the childhood skills and signals, but can also reflect other unobserved variables.
Figure 1.1: Stylized sketch of the conceptual framework that is developed in this thesis

The conceptual framework in Figure 1.1 is translated to an empirical framework in chapter four. Chapters two and three build up to this framework. In particular, chapter two only considers the effects of the childhood skills on adult offending outcomes, whereas chapter three models the interaction between offending, employment and social welfare over the adult life span. The results and methodology that are developed in these chapters are used to guide the decisions for the complete framework that is considered in chapter four.

1.3 Theoretical foundation

The conceptual framework that we sketched in Section 1.2 is based on insights from a variety of theories and perspectives from criminology, sociology, psychology and economics. In particular, the framework encompasses elements from self-control theory (Gottfredson & Hirschi, 1990), social control theory (Hirschi, 1969; Laub & Sampson, 2003) and the rational choice perspective developed in Becker (1968). Additionally, we incorporate the main insights from the recent literature on cognitive and social skill for-
mation that is developed both in economics (e.g., Heckman, 2006; Heckman, Stixrud, & Urzúa, 2006; Cunha & Heckman, 2007, 2008; Borghans, Duckworth, Heckman, & Ter Weel, 2008; Cunha, Heckman, & Schennach, 2010) and psychology (e.g., Lodi-Smith & Roberts, 2007; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007; Blonigen, 2010).

1.3.1 Criminological and economic theories

The self-control theory of Gottfredson and Hirschi (1990) argues that individual differences in offending are the result of a single higher order personality construct that is labeled self-control. A broad definition of self-control is given by the combination of personality characteristics such as impulsive behavior, conscience development, self-regulation, delay of gratification, inattention-hyperactivity, executive function, willpower and inter-temporal choice (Moffitt et al., 2011). The theory suggests that differences in self-control, at any point in time, can explain the differences in offending, and that other outcomes for aspects such as marriage and employment are essentially driven by self-control. In particular, underlying personal characteristics that select a person into anti-social behavior and delinquency, also select a person into disadvantaged labor market positions, such as welfare and unemployment, disadvantageous marital positions and other adult outcomes.

Since Gottfredson and Hirschi (1990) assume that the level of self-control is established early in life their theory can be regarded as a static theory. In Hirschi and Gottfredson (1983) they hypothesize that no combination of underlying psychological and sociological variables can explain the age-variation in offending that occurs over the life span. Only age itself explains the variation in offending that occurs with maturation. The conceptual framework of this thesis accommodates the notion that skills that are developed early in life, can explain individual-level differences in adulthood offending and other socioeconomic adult outcomes. However, as discussed in detail below, we generally allow the impacts of skills to vary with age and allow structural dynamic effects from other adult outcomes to influence offending as well. Also, we make no attempt to explicitly capture the construct of self-control, but rather investigate more interpretable childhood factors such as cognitive abilities and social skills. Arguably, self-control can be interpreted as a higher-order construct of the included
CHAPTER 1: GENERAL INTRODUCTION

childhood skills.

A second prominent perspective for offending is found in social control theory, which states that everybody is capable of offending, but the majority of individuals are restrained by the social bonds that tie them to society (Hirschi, 1969). Delinquency only becomes an option when these social ties are broken or weakened. Sampson and Laub (1993) argue that the relevant social bonds vary over the life course. For instance, social bonds with family, school and peers are important for the adolescence period. When the value that these social bonds hold for an individual exceeds the costs of offending, delinquency will become less attractive. Sampson and Laub (1993) named this value social capital, the importance that the ties to society hold for the individual. After adolescence follows a period that is characterized by exploration with limited parental control, and with the aim of establishing a unique personal identity (Arnett, 2004). In this period a shift occurs in relevant institutions of social control, from family, school and peers to more prominent bonds with partners and co-workers.

Social control theory provides a more dynamic perspective for explaining offending over the life course and we make an attempt to include aspects that represent social bonds in our empirical framework. In particular, we allow adult life outcomes and offending to mutually influence each other over the life course, as such offending can weaken social bonds that may prevent individuals from future employment and intimate relationships (Hirschi, 1969). Sampson and Laub (1997) also suggest the weakening of conventional bonds to society by the gradual process of cumulative disadvantage. Furthermore, Sampson and Laub (1993) identify the dynamic process of childhood antisocial behavior and adolescent delinquency as a possible cause of adult crime as it limits individuals from obtaining adult social bonds. Criminal behavior, and interaction with the criminal justice system, labels an individual as an offender, tainting the individual’s self-image and public identity which in turn affects future life outcomes (Nagin & Paternoster, 1991).

The concept of social capital is closely related to the concept of human capital which is prominent in economics. In particular, human capital is loosely defined as the stock of skills, both cognitive and social, that governs the “productivity” of workers (Becker, 1993). Skills can be either innate or learned. Both social capital and human capital are broadly defined concepts, which are used to explain either offending, as in Sampson and Laub (1993), or employment, as in Becker (1993). In our empirical framework we aim
to simultaneously explain employment and offending, as well as other adult outcomes, by the same large set of skills, such that social and human capital are defined by the same concepts. Of course different skills can be important when explaining offending, when compared to explaining employment.

The rational choice perspective for explaining offending is developed in Becker (1968), Ehrlich (1973) and Block and Heineke (1975). Within this framework criminal behavior is viewed as illegal employment and criminal behavior results from the risk-return trade-off between legal and illegal employment (e.g., Ehrlich, 1973; Grogger, 1998). The goal for the individual is to maximize utility, where utility can in principal be derived from a variety of situational characteristics. Some individuals choose crime rather than legitimate employment because they expect to gain more from crime, while taking into account the expected probability and severity of punishment. This perspective also implies dynamic interaction between adult outcomes and is as such reflected in the conceptual framework in Section 1.2.

Finally, the conceptual framework is related to the recent economic perspective on childhood skills formation (e.g., Heckman, Stixrud, & Urzúa, 2006; Cunha & Heckman, 2007, 2008; Cunha et al., 2010). In this perspective the formation of childhood skills is envisioned as a dynamic process in which children accumulate skills from birth endowments and parental investments. The childhood phase is thus considered to last for multiple periods and the final childhood period gives a set of skills that are used to explain socioeconomic adult outcomes (Cunha et al., 2010). The main difference with our conceptual framework is that we treat childhood as a single period. In particular, we do not model the dynamics in the formation of childhood skills, but rather observe these at the end of childhood. Instead, and of major importance, we treat adulthood as a dynamic process and allow adult outcomes to mutually influence each other over the adult life span. This allows us to assess multiplier effects from adult life transitions, which may increase or decrease the cumulative effects from the childhood skills on the subsequent adult outcomes.

1.3.2 The role of skills and personality traits

The dominant theoretical perspectives in life course criminology treat personality traits as fixed constructs that can explain level differences in offending (Nagin & Paternoster,
CHAPTER 1: GENERAL INTRODUCTION

1993). They generally do not allow for age-varying effects from different personality traits. Recent research from personality psychology has redefined personality constructs as inherently developmental such that they can explain both individual stable differences as well as differences over the life course for a variety of adult outcomes (Caspi, Roberts, & Shiner, 2005). While many personality traits are found to be consistent in their ordering of individuals in a population over time (Roberts & DelVecchio, 2000), this does not imply that their average effect on adult outcomes is stable over time.

In our empirical framework we accommodate the notion that different skills, or personality traits, can be important in explaining different adult outcomes at different stages of adulthood. For example, social skills may explain adolescent offending, whereas cognitive skills may explain adulthood employment probabilities (e.g., Heckman, Stixrud, & Urzúa, 2006; Agan, 2011; Carneiro, Crawford, & Goodman, 2011). We follow the theoretical framework developed in Blonigen (2010) where it is argued that a variety of personality traits may explain the variation that occurs in offending over the life-span. We adopt this perspective and allow different personality traits to affect adult outcomes for offending, employment, social welfare, drug use and intimate relationships in a dynamic way such that the payoffs from the traits may be different at different stages of adulthood.

1.4 Towards a testable empirical framework

The next important step that is taken in this thesis is the translation of the conceptual framework of Section 1.2 to a set of mathematical equations which summarize the empirical framework. This translation facilitates the testing of the conceptual framework using observational data and econometric methods. Ultimately, the quality of the translation determines whether we are able to provide a realistic empirical assessment of the conceptual framework. The translation requires a tradeoff between complexity and econometric feasibility. In particular, the mathematical relationships can be formulated in an excessively complex manner, which might prevent us from testing the empirical framework using observational data. On the other hand, if the translation is too simple we do not capture the essential features of the conceptual model.

The exact formulation of the empirical framework is discussed in detail in the
chapters below. Here we highlight several features of the conceptual framework and the characteristics of the observational data that need to be taken into account when translating the conceptual framework into the empirical framework. In particular, we briefly discuss the incorporation of skills, omitted variables, simultaneity, non-linearity and the presence of missing values.

The childhood skills that influence the adult outcomes are generally unobserved. Three problems exist when aiming to capture these skills. First, skills are hard to capture by a single or a few variables and it is often unclear which variables are important. This presents a variable selection problem for which many methods, such as shrinkage, bagging and pre-testing, have been proposed (Stock & Watson, 2012). Second, the variables are generally not perfectly observed. In other words, coding differences, timing issues and varying interpretations of different psychologists may impose the presence of measurement error. Third, even when variable selection and measurement error are taken into account missing values are often present. Since typically not all measurements are collected for all individuals, and changes in the types of measurements collected over time, may occur, the presence of missing values is quite common in measurements for childhood skills.

To overcome these issues the empirical framework that is adopted in this thesis proposes two different methods. The first method, which is used in chapter two, uses a variety of personality traits that are assessed by psychologists and translated into norm values to approximate the childhood skills. Data imputation methods are then used to accommodate missing values. This approach has the benefit that the payoffs from the skills are easy to interpret, but the downside is that measurement error remains present. The second method which is adopted in the fourth chapter follows Heckman, Stixrud, and Urzúa (2006) and defines the skills as latent factors that are estimated using a linear factor model which incorporates a large number of measurements for the skills. This approach has the advantage that measurement error is taken into account, but makes the factors more difficult to interpret.

The stylized conceptual framework in Figure 1.1 does not include all aspects of childhood and adulthood that might be relevant for predicting the adult outcomes. For example, we do not model adult outcomes for living situations, which may affect the other adult outcomes. This implies that the translation of the conceptual framework to the empirical framework needs to incorporate unobserved, or omitted,
variables. Statistically speaking, the presence of omitted variables creates spurious correlation between the subsequent adult outcome variables. To capture spurious correlation we allow for serial correlation in the error terms of the empirical framework. In particular, we either model the error terms using autoregressive processes (Keane, 1994), or we incorporate a dynamic factor structure for the error terms (Pesaran, 2006). Both approaches serve to reduce the influence of omitted variables on the structural relationships that are of interest in the empirical framework.

The simultaneity problem, or reverse causality problem as discussed by Ehrlich (1973), arises when the observed correlation between the adult outcome variables stems from bi-directional causal relationships. For example, offending affects employment and employment also affects offending. The consequence of the simultaneity problem is that the identification of causal relationships is complicated. Valid instrumental variables are generally not available for individual-level studies and the sample sizes are too small to achieve identification from exogenous shocks. Instead, we take a time series approach and are mainly interested in whether current outcomes can predict future outcomes (Granger, 1980). This form of Granger-causality is reflected in Figure 1.1 by the fact that there are no arrows between the different adult outcomes within each adulthood period.

Next to the aforementioned problems of skills measurements, omitted variables and simultaneity, the individual-level setting presents a number of additional modeling difficulties that need to be addressed. First, the adult outcomes outcomes are typically non-linear, or non-Gaussian, in nature. For example, offending outcomes are often coded as binary variables, where the value one indicates that an offense is committed. Second, missing observations are common as many individuals are only partly observed. For example, incarceration spells cause individuals to temporarily lose their ability to make decisions during adulthood. These difficulties do not pose difficulties for translating the conceptual framework into an empirical framework, but they do make testing the empirical framework using econometric methods complicated.

This thesis translates the conceptual framework into an empirical framework while taking into account the translation difficulties. Such a theory-based modeling approach requires assumptions and we state these explicitly throughout the chapters. More discussion for the general modeling approach can be found in Keane (2010) and Wolpin (2013).
1.5 Data summary

1.5.1 17up study

In chapters two, three and four of this thesis we use data from the NSCR 17up study. This is a longitudinal study that follows disadvantaged youths into adulthood. The original 270 males and 270 females were all discharged from a juvenile treatment facility in the 1990s. We included all individuals who stayed in the institution for at least three months and were not being treated for sexual offending behavior\(^4\). Three main sources of data are used for the analyses: treatment files from the institution, official register data, and retrospective interview data for a subsample of the original group.

**Treatment files**

Measurements that are related to childhood skills are obtained from the individuals’ treatment files. Permission for the use of the treatment files was obtained from the Department of Corrections (DJI). The files were obtained from the archive of the juvenile treatment facility. These files generally contained the results from psychological and psychiatric tests, advisory notes on extensions, and treatment evaluations. The reports in the files are in most cases prepared by forensic psychologists and psychiatrists. However, also external reports from the Dutch Child Protection Agency and other organizations responsible for the supervision of juveniles were found. For all individuals progress reports regarding their treatment had been compiled by a multi-disciplinary team.

Although the contents of the files varied between individuals, we were able to extract a large number of common items. These measurements are able to capture broad measures for both cognitive and social skills. In chapter two we use specific items from the files that capture personality traits, which can be seen as proxies for the cognitive and social skills. In chapter four we follow Heckman, Stixrud, and Urzúa (2006) and use a factor model to convert large vectors of measurements into low-dimensional factor scores for cognitive and social skills. Both approaches have their own benefits

\(^4\)We excluded the sex offenders as there are reasons to believe that their behavioral problems and treatment are different from non-sexual offenders, see for example the meta-analysis of Seto and Lalumière (2010).
and drawbacks. In particular, by selecting specific personality constructs the interpretation for the effects is simple. A drawback is that specific measurements are often contaminated with measurement error. The factor model approach that is adopted in chapter four circumvents this by using a factor model that incorporates measurement error. The drawback of the factor model approach is that the interpretation for the effects from the factor scores is less straightforward.

Official register data

In this thesis we make use of three kinds of register data. First, the offending and incarceration data are obtained from convictions registered in the Judicial Documentation (JD) abstracts of The Netherlands Ministry of Justice. Permission for the use of the criminal records was obtained from the Netherlands Ministry of Justice (JustID). These are comparable to rap sheets in the US. The abstracts contain information on every case that is sent to the Public Prosecutor’s Office and the decision that follows on it. They also contain information on date and type of the offense. The abstracts are available for each individual from age 12 and onwards, 12 being the age of criminal responsibility.

We generally consider three categories for offending that include serious, property and violent offenses. The serious offenses are constructed following the definition given in Loeber, Farrington, and Washburn (1998) which includes all violent offenses, felony larceny, auto theft, burglary, breaking and entering, carjacking, forgery and counterfeiting, fraud, dealing in stolen property, embezzlement, drug trafficking, arson, weapons violations and firearms violations. The property offending category includes crimes such as embezzlement, theft, forgery and counterfeiting, breaking and entering, burglary, fraud and dealing in stolen property. The violent offending category includes assaults, threats, homicides, sexual offenses, robberies and kidnapping.

Second, we use official register data to construct the employment and social welfare outcomes. This is obtained between 2007 and 2009 from the Ministry of Social Affairs and Employment (SoZaWe). The information consists of individual-level employment and social welfare histories from 1992 onwards. For each employment spell we know the exact start and ending date of the contract. Whether a position was full-time or part-time remains unknown to us as we have no information on the exact amount of
hours spent working. Also, at the time of collecting the SoZaWe database contained no reliable information on wages.

We consider three different employment variables, which include spells that pertain to any type of employment, spells that pertain to regular employment and spells that pertain to employment via a temporary job agency. The latter distinction allow us to investigate the effect of job stability, since employment through a temporary job agency in the Netherlands often is seasonal or project based, and generally lasts for short spells of a few weeks to a few months providing little long-term prospects compared with regular jobs.

Additionally, we consider three forms of social welfare: unemployment insurance, disability insurance and public assistance. Unemployment insurance consists of payments for those who have lost their job, whereas disability insurance provides payments during illness. Insurance policies are temporary in nature. When individuals do not manage to find employment within the designated period or remain unable to work, public assistance is available to replace income. Public assistance, the most important welfare policy, is meant to assure recipients a minimum income needed for subsistence. Such benefits do not require proof of anything other than financial need, nor are they conditional on prior employment. More discussion for the requirements and details for social welfare are given in chapter three.

Third we use municipal registries to construct variables for marriage, divorce and the number of children. Permission for the use of municipal records was obtained from the Ministry of the Interior and Kingdom Relations (BZK). While our central focus is not on these variables we often include these variables as controls to limit their influence on the relationships among other adult outcomes variables.

Retrospective interview data

Face-to-face interviews were conducted with a subsample of the original 540 males and females. At the start of the interview phase in 2010, 22 of the original individuals had died, 14 had emigrated, 5 were living in institutions and another 19 could not be traced. The remaining 499 males and females were approached for interviews. Out of the 499 individuals 116 males and 132 females completed a full interview, after giving informed consent. Permission for the follow up study was given by the Department of
CHAPTER 1: GENERAL INTRODUCTION

Correction (DJI) and the ethics committee CERCO of the VU University Amsterdam. A response analysis was conducted to verify the representativeness of the interviewed versus the original sample. Besides individuals without a regular place of living the subsample can be regarded as representative for the original sample. Further details for the interviews and the response analysis are given in van der Geest, Bijleveld, and Verbruggen (2013).

The interview consisted of two parts from which we obtained data. First, a structured questionnaire was filled in which covered a wide array of topics. For the purpose of this thesis we only used the questions related to education. The second part of the interview included filling in a life-history calendar. This tool aims to retrospectively reconstruct the life of the individual for a variety of adult life domains (Caspi, Moffitt, Thornton, & Freedman, 1996). For the purpose of this study we used the information related to drug use and intimate relationships.

1.5.2 TRANS-5 study

In chapter five of this dissertation we make use of data from the TRANS-5 study, see Bijleveld, Wijkman, and Stuifbergen (2007). The TRANS-5 dataset consists of observations for five generations of families (G1-G5). The five generations are ancestors and descendants of the original sample, which consists of 198 males who were institutionalized in a reform school between 1911 and 1914. While, the behavioral problems of these youths may not be identical to the behavioral problems of the youths from the 17-up study that were released from a juvenile treatment facility in the 1990s, they can both be regarded as disadvantaged youths, albeit in their own time period. The offspring of the G2 males and their spouses were traced in Dutch genealogical and municipal records, resulting in a 100% retrieval rate. Sample members of those G2 who had emigrated, or died before the age of 21, were considered lost to follow-up and their descendants were not traced. After removing these, 181 men remained who had offspring that is labeled G3; subsequent generations are labeled as G4 and G5.

In chapter five we rely on four generations, G2-G5, as criminal data for the G1 generation was available but was probably incomplete. In total, we included 4,120 individuals, to which an additional 1,919 spouses could be linked, resulting in 6,039 men and women. Each individual is included from age 12 to 60. Observations between
1920 and 2005 are included for generations G2 to G5.

1.6 Summary of the remainder

The remainder of this thesis is organized as follows. Chapter two examines to what extent cognitive and social skills that are formed during childhood and adolescence can explain the age-crime curve for disadvantaged youths. We propose a new decomposition model for explaining adolescent and adulthood offending, which incorporates a dynamic factor structure that captures age-varying payoffs from cognitive and social skills. We test whether personality traits can explain individual and age-varying differences in offending.

Chapter three tests economic and sociological theories for the relationship between employment and crime, where social welfare is used as an identifying mechanism. Childhood is treated as an unobserved period for which we statistically control. We simultaneously model the offending, employment and social welfare variables using a dynamic discrete choice model, where we allow for state dependence, reciprocal effects and time-varying unobserved heterogeneity. The chapter documents the interaction between offending, employment and social welfare during adulthood.

Chapter four combines the insights from chapters two and three and develops a complete framework that decomposes multiple adult outcomes into effects from childhood skills, signals and adult life transitions. We consider both cognitive and social skills, which may partially determine childhood signals for education and criminal records. Both the skills and the signals, are subsequently allowed to affect adult outcomes for offending, employment, social welfare, drug use and intimate relationships.

Chapter five provides a historical perspective for the relationship between unemployment and crime. We study to what extent the causal relationship between unemployment and crime has changed throughout the last century for families of disadvantaged youths.

Chapter six summarizes the main empirical findings and discusses the implications of the findings for criminological life course theories and public policy. We finish by discussing some directions for future research.
Chapter 2

Explaining the Age-Crime Curve of Disadvantaged Youths by Dynamic Payoffs from Cognitive and Social Skills

2.1 Introduction

Arguably the most firmly established empirical finding in life course criminology is the age-crime curve (Nagin & Land, 1993). This curve, which already appeared in Quetelet (1831), documents that criminal behavior starts in late childhood, peaks during adolescence (age 15-19) and decreases during adulthood (e.g., Farrington, 1986; Farrington, Piquero, & Jennings, 2013). While similar curves are observed across cultures, mild variation exists in the shape of the curve for different time periods and crime types (Steffensmeier, Andersen, & Harer, 1989). Despite the empirical robustness of the age-crime curve, little consensus exists about its etiology (Osgood, 2005, 2012). Some scholars have suggested that age itself explains the shape of the age-crime curve and that no underlying sociological or psychological variables explain the curve (Hirschi & Gottfredson, 1983; Gottfredson & Hirschi, 1990). Other more dynamic perspectives suggest that the weakening or strengthening of social bonds, such as those from school, employment and marriage, that tie individuals to society, shape the age-crime curve.
A second literature, which combines insights from psychology and economics, has established that broad measures of childhood skills, including cognitive and social skills, have lasting impacts on socioeconomic adult outcomes, see Heckman (2006), Almond and Currie (2011) and Heckman and Kautz (2013) for recent reviews. Childhood skills are accumulated in a dynamic process where parental investments and initial birth endowments, which may depend on parental skills, are the inputs (Cunha & Heckman, 2007). By the end of childhood these skills have become important predictors for a broad array of adult outcomes (Heckman, Stixrud, & Urzúa, 2006; Cunha & Heckman, 2008; Cunha et al., 2010). Criminal behavior is often among the socioeconomic outcomes that are studied, see for examples Heckman, Stixrud, and Urzúa (2006), Agan (2011) and Carneiro et al. (2011). These contributions show that social skills, in addition to cognitive ability, are important predictors for adult offending. This is underlined by the fact that interventions aimed at improving social skills have shown to be effective in reducing crime (Hill et al., 2011).

In this chapter we empirically study to what extent skills that are obtained during childhood and adolescence can explain the age-crime curve. More specifically, we investigate whether age-varying “payoffs” from childhood and adolescent skills can explain the variation in crime that occurs with age (Blonigen, 2010). To do so, we develop a novel statistical model that incorporates childhood and adolescent skills in a dynamic model for adolescent and adulthood offending. We include cognitive skills and a variety of social skills, amongst which personality traits such as neuroticism, extroversion, impulsiveness, thrill seeking and conscience development, which arguably co-determine social skills1, and allow these to affect the age-crime curve according to an age-varying payoff vector. In doing so we preserve the property of rank-order stability for the skills but accommodate the feature that rank-order stability does not imply absolute stability (Blonigen, 2010)2. We explicitly model the intuitively appealing notions that (a) multiple skills can be important in determining the age-crime curve and (b) different

---

1 Different labels and definitions for social skills can be chosen. For example, social skills are also often referred to as personality traits or non-cognitive skills (Borghans et al., 2008). In this paper we follow Cunha and Heckman (2007) and use the term skills which suggests a dynamic process of formation in which skills are formed and can possibly be improved over the life course.

2 Rank-order stability refers to consistency in the relative ordering of individuals’ skills in a population over time (Blonigen, 2010).
skills can be important in explaining the age-crime curve at different stages in life (Cunha & Heckman, 2007; Blonigen, 2010). We point out that the model does not aim to highlight different types of criminal careers in the sense of Moffitt (1993) and Nagin (2005), but rather investigates whether the heterogeneity in the offending outcomes, between individuals and with age, can be explained by differences that stem from the skills that are obtained during childhood and adolescence. Ex post analysis of the model residuals can reveal trajectories that are unexplained by the childhood skills. Such unexplained trajectories would motivate further augmentation of the model to incorporate other childhood factors and adult life events, such as employment and marriage, which are deemed important in social control theory (Hirschi, 1969; Laub & Sampson, 2003).

From a statistical perspective the model can be viewed as a nonlinear state space model (Cappé, Moulines, & Rydén, 2005; Durbin & Koopman, 2012), where the state space is formed by a dynamic factor structure that incorporates the childhood skills. The dynamic factor structure incorporates multiple skills with age-varying payoffs. We suggest a likelihood-based estimation method for the fixed model parameters which is based on the importance sampling methods developed in Jungbacker and Koopman (2007). Once the model parameters are estimated the model residuals can be extracted and further analyzed.

To illustrate the model we consider samples of disadvantaged males and females who were institutionalized in a juvenile treatment facility in the 1990s in the Netherlands. The sample members are at high risk of offending (van der Geest, Blokland, & Bijleveld, 2009, 2011). During their stay in the treatment facility extensive files were constructed by psychologists and other caretakers. From these files we obtained measurements that are used to approximate the cognitive and social skills. This information is supplemented with official criminal records from which we construct the adult offending variables. By result, we are able to study the effects of childhood skills on serious, property and violent offending for both males and females.

The remainder of this chapter is organized as follows. In the next section we discuss life course criminological theories that aim to explain the age-crime curve. Section 2.3 outlines our research goals. In Section 2.4 we present our statistical model that serves to decompose the individual offending trajectories into payoffs from cognitive and social skills. In Section 2.5 we discuss the sample for which we have obtained data. Here we
also detail the construction of the variables for the skills and the offending outcomes. The results are shown in Section 2.6 and Section 2.7 discusses the implications of our findings.

2.2 Explaining the age-crime curve

We review the main life course criminological theories that aim to explain the age-crime curve. We follow Blokland and Nieuwbeerta (2010) and consider the self-control theory of Gottfredson and Hirschi (1990), the age-graded theory of social control of Sampson and Laub (1993) and Laub and Sampson (2003), and the dual-taxonomy of Moffitt (1993), as the most important life course theories. We concentrate our discussion around the roles that these theories assign to cognitive and social skills. In general, the view that individual differences in skills only explain average age-invariant differences in offending is widespread in life course criminology and is as such reflected in the dominant theoretical perspectives (Nagin & Paternoster, 1993; Blonigen, 2010).

The self-control theory of Gottfredson and Hirschi (1990) argues that individual differences in offending are the result of a single personality construct labeled self-control\(^3\). They suggest that differences in self-control are established early in life and represent a stable trait over the life course. As such they rule out that age-varying effects from self-control are a possible explanation for desistance from crime in adulthood (Blonigen, 2010). Instead they argue that age itself is the main explanation for the shape of the age-curve and that no combination of psychological and sociological variables can replicate the curve (Hirschi & Gottfredson, 1983; Sweeten, Piquero, & Steinberg, 2013).

The age-graded theory of social control poses a more dynamic explanation for the age-crime curve (Laub & Sampson, 2003). It argues that criminal involvement results from a lack of informal social controls and that controls vary with age. Transitions in life domains, such as marriage and employment, can act as turning points that result in desistance from crime. Laub and Sampson (2003) argue that such transitions are to

\(^3\)According to Moffitt et al. (2011) self-control is interpretable as a higher order construct of more clearly defined personality traits such as impulsive behavior, conscience development, self-regulation, delay of gratification, inattention-hyperactivity, executive function, willpower and inter-temporal choice.
CHAPTER 2: EXPLAINING THE AGE-CRIME CURVE

a certain extent chance events occurring independently of underlying traits and skills.

Intuitively, both underlying skills and ability, as well as life course transitions that occur by chance, are likely to explain the age-crime curve to some extent (Blokland & Nieuwbeerta, 2010). Sweeten et al. (2013) show that a combination of variables that proxy social control, learning, strain, psychosocial maturity and rational choice can explain a significant part of the relationship between age and crime. Also, Loeber et al. (2012) find that impulsiveness and intelligence significantly predict the age-crime curve for a sample of boys from the Pittsburgh Youth Study.

A different perspective for the building blocks of the age-curve is found in taxonomic theories. These theories can be seen as complementary to self-control and social control theories (Sampson & Laub, 2005). In particular, they suggest that the aggregate age-crime curve can be decomposed into multiple underlying trajectories which typically summarize the criminal behavior of subgroups, that are labeled life course persistent and adolescent limited offenders (e.g., Moffitt, 1993; Moffitt & Caspi, 2001). The life course persistent offenders show continuing anti-social behavior which starts in late childhood and lasts well into adulthood. Adolescent limited offenders have much shorter criminal careers which are typically restricted to the adolescent and early adulthood years. This suggests that there exists substantial heterogeneity around the age-crime curve in the form of potential subgroups of offending individuals. Consistent with Moffitt’s typology several empirical studies have shown that these subgroups have been shown to have different personality characteristics (e.g., van der Geest et al., 2009).

Summarizing, the main life course theories assume that skills are formed early in life and they are treated as fixed constructs. In other words, they can explain level differences in offending but not changes in offending over the life course. Also, substantial heterogeneity in the age-crime curve exists within a certain population. While the shape of the average age-crime curve is broadly accepted, different underlying subgroups can have vastly different curves. Finally, several possible adulthood determinants for the age-crime curve have been hypothesized. Often these determinants are assumed to act independently from the skills of the population members (Laub & Sampson, 2003).

---

4Several studies have documented more than two subgroups, see the discussion and references in Nagin (2005).

25
2.3 The current study

In the current study we test a dynamic perspective for childhood skills. In particular, recent research from personality psychology has redefined personality constructs as inherently developmental such that they can explain both individual stable differences as well as differences over the life course (Caspi et al., 2005). While many personality traits are found to consistently order individuals in a population over time (Roberts & DelVecchio, 2000), this does not imply that their average effect, or payoff, on the probability for offending is stable over time (Blonigen, 2010). In this paper we test whether such mean-level changes can explain the variation in the level of the age-crime curve within a population of disadvantaged youths. We assume rank-order stability and do not investigate individual-level change.

We propose a statistical model that posits a dynamic factor structure for the skills. As such different skills can increase, or decrease, the propensity towards criminal behavior at different stages in life. Further, we consider different crime types since it is likely that different skills impact different types of crime. Throughout our study we include control variables for employment, marriage and incarceration. We compare the payoffs from the childhood skills between models with and without control variables.

2.4 Statistical model

When a particular segment of the population enters early adulthood we assume that differences in multiple skills have already been formed (e.g., Cunha & Heckman, 2007).

For our model we summarize the resulting childhood and adolescent skills for individual \( i \) in the \( r \times 1 \) vector \( f_i \), which includes both cognitive and social skills. Given \( f_i \) individual \( i \) enters the early adulthood period. During adulthood, which lasts for \( T \).

---

5For example, evidence suggests that IQ scores become stable from age 10 onwards (Hopkins & Brecht, 1975). Also, the development of the prefrontal cortex that governs emotions and self-control is completed in the early 20’s (Dahl, 2004).

6In our current model we consider the skills as age-invariant. They are the result from the childhood and adolescent periods. A more elaborate model would model the dynamics in the accumulation of such skills during childhood, see for example Cunha et al. (2010).

7In our empirical application the skills considered are intelligence, neuroticism, impulsiveness, extroversion, thrill seeking and conscience development.

8In our empirical application we consider this point around age 17, which is the average age that the studied individuals left the treatment facility.
years with each year indexed by $t$, the individual has the opportunity to offend. Let $y_{i,t}$ denote a binary indicator, which is equal to one if and only if individual $i$ commits an offense at age $t$. Different types of offending can be considered. A binary logistic model for the offending variables is given by

$$y_{i,t} \sim \text{Binary}(\pi_{i,t}), \quad \theta_{i,t} = \log \frac{\pi_{i,t}}{1 - \pi_{i,t}}, \quad (2.1)$$

where $\pi_{i,t}$ is the probability that individual $i$ commits an offense at age $t$ and $\theta_{i,t}$ is the transformed probability, where the transformation is performed by the canonical link function for binary models, see Cox and Snell (1989). When averaging $\pi_{i,t}$ over the individuals a probabilistic age-crime curve can be obtained. As such $\pi_{i,1}, \ldots, \pi_{i,T}$ are interpretable as individual-specific probabilistic age-crime curves.

The transformed probabilities $\theta_{i,t}$ are given by the linear decomposition model:

$$\theta_{i,t} = c + f_i'\delta_t + X_{i,t}'\beta_t + e_{i,t}, \quad (2.2)$$

where $c$ is the constant, $\delta_t$ determines the payoff from the childhood skills $f_i$, $X_{i,t}$ is the $k \times 1$ vector of observed control variables that is measured by $\beta_t$ and $e_{i,t}$ is the disturbance term. The payoffs $\delta_t$ are allowed to vary with age to accommodate the feature that the payoffs from childhood and adolescent skills can vary with age while preserving their rank-stability (Blonigen, 2010). The vector $X_{i,t}$ includes variables that can influence criminal behavior during adulthood. In our empirical application we include variables for employment, marriage and incarceration. We emphasize that the effects of these variables can also vary with age. For example, employment is often found to increase delinquency for adolescents, but reduces criminal behavior for adults (e.g., Shover, 1996; Uggen, 2000; Paternoster, Bushway, Brame, & Apel, 2003). We model the elements of the payoff vectors $\delta_t$ and $\beta_t$ using cubic spline functions. This is justified when we assume that the payoffs vary smoothly with age.

We model the error term as a stationary autoregressive process of order one. In particular the model for $e_{i,t}$ is given by

$$e_{i,t} = \gamma e_{i,t-1} + \sqrt{1 - \gamma^2}\eta_{i,t}, \quad \eta_{i,t} \sim N(0, 1), \quad (2.3)$$

\footnote{A model that would not impose rank-stability for the skills would replace $f_i$ by $f_{i,t}$. This however requires multiple measurements for the skills over time.}
where $\gamma$ is the autoregressive coefficient and $\eta_{i,t}$ is the disturbance term which follows a standard normal distribution. By scaling $\eta_{i,t}$ by $\sqrt{1-\gamma^2}$ the process for $e_{i,t}$ is normalized such that $\text{Var}(e_{i,t}) = 1$, see also Heiss (2008).

The error term $e_{i,t}$ plays an important role in our model. First, suppose that $\delta_t = 0$ and $\beta_t = 0$ for $t = 1, \ldots, T$. Then the cross-sectional averages of the estimated error terms $N^{-1} \sum_{i=1}^{N} \hat{e}_{i,t}$ capture the age-crime curve. When $\delta_t \neq 0$ and $\beta_t \neq 0$ the cross-sectional averages of the estimated error terms capture that part of the age-crime curve that is unexplained by the payoffs from the skills and the control variables. Thus, when comparing these models we can determine which part of the age-crime curve is explained by the childhood skills and the control variables. An alternative, but equivalent, interpretation is that the error terms capture spurious persistence in the adult offending outcomes (Keane, 1994). In other words, many shocks may affect the offending probability and most of these are unobserved. The error term captures the shocks that are unexplained by the skills and observed control variables.

The fixed model parameters are estimated using the Monte Carlo maximum likelihood methods developed in Jungbacker and Koopman (2007). A textbook treatment for the estimation method is provided in Durbin and Koopman (2012, Part 2).

### 2.5 Data description

The sample that we consider consists of 270 males and 270 females who were institutionalized in a juvenile treatment facility in the early 1990s in The Netherlands. We combine information from their treatment files with official register data. The treatment files provide the measurements for the cognitive and social skills, and the register data provides the information for the adulthood offending outcomes.

In the Netherlands juveniles can be sent to treatment facilities for various reasons. Examples include criminal behavior, disrupted family environments and serious behavioral problems. Wijkman, van der Geest, and Bijleveld (2006) show that civil or criminal law reasons for treatment do not explain differences in behavioral problems. The ages in the facility range between 10 and 20 over the whole period, with the median around 16-17. While staying in the facility low-level secondary and vocational education is provided and behavioral problems are treated. In the next two subsections we describe the variables that were used to proxy the cognitive and social skills, and the
2.5.1 Cognitive and social skill variables

We obtained a number of measurements for the cognitive and social skills from the individuals' treatment files. These files were obtained from the archive of the juvenile treatment facility. They generally included the summaries from psychological and psychiatric tests, advisory notes on extensions, and treatment evaluations. The reports in the files were prepared by forensic psychologists, psychiatrists and other caretakers from the treatment facility. In addition, also external reports from the Dutch Child Protection Agency and other organizations responsible for the supervision of juveniles are sometimes included. For all individuals, progress reports regarding their treatment were compiled by a multi-disciplinary team.

Although the contents of the files varied between individuals, we were able to extract a number of common items. In particular, we include variables for intelligence, neuroticism, impulsiveness, extroversion, thrill seeking and conscience development. The descriptive statistics for the variables are given in Table 2.1. The intelligence scores had been measured using the Wechsler Intelligence Scale for Children (revised for The Netherlands) and the Raven Progressive Matrices. These were then turned into categories according to prevailing norm values that ranged between very low and gifted. The females score slightly higher on the intelligence scale, when compared to the males.

For measuring social skills we employ measurements for a large number of personality traits. The variables for neuroticism, impulsiveness, extroversion and thrill seeking were derived from standard questionnaires used in the files and validated self reports. Relevant guidelines, such as the ATL (Adolescent Temperament List) and the NPV-J (Netherlands Personality Questionnaire - Youth) were used to construct the norm values. The level of conscience was measured by the researchers based on the information available in the files. We refer to van der Geest and Bijleveld (2008) for additional discussion for the construction of these measurements. The descriptive statistics indicate that females scored higher on all skills except for conscience development. To avoid biases from systematic differences between adolescent males and females we treat these samples separately.
Table 2.1: Summary statistics for measurements for cognitive and social skills.

<table>
<thead>
<tr>
<th></th>
<th>Males N = 270</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Intelligence</td>
<td>2.705</td>
<td>0.841</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>3.167</td>
<td>0.797</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Impulsiveness</td>
<td>2.965</td>
<td>0.847</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Extroversion</td>
<td>1.608</td>
<td>0.752</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Thrill seeking</td>
<td>2.573</td>
<td>0.763</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Conscience</td>
<td>2.180</td>
<td>0.741</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Females N = 270</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Intelligence</td>
<td>2.784</td>
<td>0.916</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>3.200</td>
<td>0.892</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Impulsiveness</td>
<td>3.099</td>
<td>0.887</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Extroversion</td>
<td>2.175</td>
<td>0.965</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Thrill seeking</td>
<td>3.306</td>
<td>1.192</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Conscience</td>
<td>1.753</td>
<td>0.736</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

2.5.2 Criminal records data

The offending data is obtained from registered judicial documentation, which contains abstracts of the Netherlands Ministry of Justice. These are comparable to rap sheets in the US. The abstracts contain information on every case that is sent to the Public Prosecutor’s Office and the verdict that follows from it. They also contain information on the date and the type of the offense. The abstracts are available for each individual from age 12 and onwards, 12 being the age of criminal responsibility. We excluded cases were the verdict was an acquittal or a dismissal for lack of evidence.

We include outcome variables for serious, property and violent offenses. The serious offenses are constructed following the definition given in Loeber et al. (1998) and the violent and property offense categories are largely subgroups of the serious offense category. Binary outcome variables are constructed to be equal to one whenever at least one offense was committed in a specific age period.

The offending rates are presented in Figure 2.1. These rates only include individuals after release dismissing ages that they were incarcerated. On average the rates are

\[^{10}\text{This classification includes all violent offenses, felony larceny, auto theft, burglary, breaking and entering, carjacking, forgery and counterfeiting, fraud, dealing in stolen property, embezzlement, drug trafficking, arson, weapons violation and firearms relations}\]
higher for males when compared to the females. The majority of the serious offenses are property offenses, which are declining gradually with age for males. The male violent offending rate is lower on average but declines slower with maturation. For females we find a strong drop in all offending rates after age 18. From age 20 onwards this rate picks up again. The strong drop may be related to the fact that many women had children at this age; see the discussion in Zoutewelle-Terovan, van der Geest, Bijleveld, and Liefboer (2014).

2.5.3 Control variables

We include adulthood control variables for employment, marriage and incarceration. An individual-level employment variable was constructed to include the number of months per year that the individual was employed. This information was obtained from the database “SUWINET” of the Ministry of Social Affairs and Employment.
in the Netherlands. The database holds individual-level information on employment by an employer and by a temporary employment agency. We included all months in which an individual was continuously employed. Similarly, we include variables that indicate the number of months that an individual was married and incarcerated. These variables are all dynamic variables and may be associated with changes in offending. By including employment, marriage and incarceration we control for their influence on the relationship between offending and the childhood skills.

2.6 Results

In this section we present the empirical results for the samples of disadvantaged males and females that are discussed in Section 2.5. The parameters of the statistical model of Section 2.4 are estimated and interpreted. We divide the discussion of the results in two parts. In Section 2.6.1 we discuss the payoffs from the childhood skills on adult offending outcomes. In Section 2.6.2 we study to what extent the childhood skills explain the aggregate age-crime curves.

2.6.1 Payoffs from childhood skills

The estimated payoffs for the childhood and adolescent skills for males are summarized in Figure 2.2. We show the maximum likelihood estimates for the vector $\delta_t$ for serious, property and violent offending. Two estimates are shown: one for the model with observed control variables for employment, marriage and incarceration, and one for the model without control variables. Additionally, we show the 95% confidence bounds for the estimates for the model with control variables.

We find that substantial variation exists in the payoffs over the adult life span. Our measure for cognitive skills (intelligence) shows that differences in this skill do not explain much variation in offending until age 25. After 25 the variation in cognitive skills significantly explain serious and property offending. In particular, higher cognitive skills imply lower offending probabilities. For violent offending we find the same pattern but the estimates are not significantly different from zero. Further, we find that the differences in the payoffs between the models with and without control variables are small. Thus, adding additional adulthood variables does not significantly
alter the payoffs from the childhood skills. This holds even though we also allowed the
effects of the control variables to vary with age.

The payoffs from social skills, that include measurements for neuroticism, impulsiveness, extroversion, thrill seeking and conscience development, also vary with age. The payoff from neuroticism, which captures the tendency to worry, indicates that differences in this skill significantly increase offending for age 16. After this the estimates decrease sharply and do not differ significantly from zero anymore. This implies that neuroticism is only associated with adolescent offending in our sample. The payoffs for impulsiveness are never significant. Our measure for extroversion, which captures the propensity towards sociability, significantly reduces serious offending between ages 16 and 20. For violent offending this estimate is only significant for ages 16 and 17, and it is never significant for property offending. The estimates for thrill seeking are significant between ages 16 and 18 for serious and violent offending. During these periods thrill seeking increase the probability for offending. The estimates for conscience development are significant between ages 21 and 27. During this period the differences in conscience development imply that more developed individuals have a lower propensity towards offending. These estimates are significant for serious and property offending and not for violent offending.

Summarizing, for males we conclude that cognitive skills are important in reducing offending after the adolescent period. Differences in social skills explain offending mainly during the adolescent period. Overall the payoffs for models with and without control variables are small and the patterns for different types of crime are similar.

For females the estimated payoffs from childhood skills are shown in Figure 2.3. The lines are defined similar as for males. We find that differences in cognitive skills, as measured by intelligence are less associated with reductions in offending when compared to males. While intelligence has an initial reducing effect for offending (not significant) the payoff is above zero from 26 onwards and becomes almost significant at the end of the observational period. The slope of the payoff vector is thus reversed when compared to the males. This holds for all offending categories considered.

The social skills that are captured by neuroticism and impulsiveness show similar payoff patterns as for males. Neuroticism is more associated with adolescent offending, albeit not significant for females. The payoff from this skill decreases with age. Impulsiveness is more important for females in explaining offending at the end of adult-
hood. After age 26 this coefficient becomes negative and after age 30 the estimate is significant. Extroversion shows that differences in sociability are important in explaining offending during adolescence. In particular, higher scores on extroversion imply a higher offending probability during adolescence. This is the opposite from what was found for males. The payoffs from thrill seeking are not significant for serious and property offending. However, thrill seeking does significantly increase the violent offending probability between ages 16 and 19. The payoffs from conscience development are not significant for females.

Summarizing, for females cognitive skills seems less important in explaining offending behavior. Social skills for impulsiveness, extroversion and thrill seeking explain more variation. However, even the payoffs from these skills are often not significant. Interestingly, for females we find different patterns for the payoffs between property and violent offending. The differences between the models with controls and without are even smaller for females when compared to males.

For all estimated models the control variables have their expected signs and patterns. For both males and females employment and marriage reduce offending from ages 21 and 25 onwards, respectively, and incarceration increases the offending probability over the entire observational period. Since these control variables are not of major interest in our study we do not discuss them further.

For most estimates the 95% confidence bounds are wide. This is mainly the result of the small sample sizes. Additionally, we are considering samples of disadvantaged youths for whom the differences in skills are not very large. Nevertheless, even for these small samples the payoffs clearly indicate that different skills are important at different stages in life and that the effects vary over the life span.

2.6.2 Residual model diagnostics

We compare the residuals between the models with and without childhood skills. We only consider models without control variables such that we can clearly illustrate the mean-level age-varying influence of the differences in skills on the average age-crime curve. Also, since the payoffs from the models with and without control variables are not very different little differences in the aggregate residuals were found between models with and without control variables.
For males the average model residuals are shown in Figure 2.4. We find that when
the payoffs from the childhood skills are not included the average posterior mean resid-
uals mimic, up to a level shift, the empirical age-crime curves in Figure 2.1. When the
payoffs from the childhood skills are taken into account the patterns nearly completely
disappear. The average posterior means of the residuals are now much flatter and
do not show the distinct peak during the adolescent period. This holds for serious,
property and violent offending. Thus, when we allow the payoffs from the skills to vary
with age we are able to largely explain the shapes of the average age-crime curves.

For females the results are shown in Figure 2.5. We find similar results. When
excluding the dynamic payoffs from the skills the average residuals mimic the age-crime

---

**Figure 2.2:** Age-varying payoffs for males from childhood skills. The rows indicate:
(i) serious offending, (ii) property offending and (iii) violent offending. The bold lines
are the estimated payoffs from the model that includes employment, marriage and
detention as control variables. The thin dotted lines are the 95% confidence bounds
for this estimate. The thick dotted lines are the estimated payoffs from the model
without observed control variables.
curve. And when the skills are included the shape of the aggregated age-crime curves is largely explained. For violent offending the age-crime curve is not well identified, even without the skills. When including the skills for violent offending we find only a mean difference for the posterior residuals but little changes in the shape of the curve, which was already flat.

2.7 Discussion and conclusion

In this chapter we proposed a new decomposition model for the age-crime curve based on state space methodology. The model empirically separates the influences of child-
hood and adolescent skills from other adult life influences. The main novelty in our framework is that we allowed the effects of the childhood skills to vary with age. This provided us with a direct test for the dynamic influences of skills on the age-crime curve.

We tested the model using data for samples of disadvantaged males and females. All sample members were institutionalized in their teenage years and come from problematic backgrounds. We studied the explanatory power of their childhood skills on serious, property and violent adulthood offending. From the empirical results we are able to draw four broad conclusions.

First, the payoffs from childhood skills on offending outcomes are not stable over the life span. In accordance with the recent literature on personality development we find that the payoffs from the differences in the included skills change with age (Caspi et al., 2005; Blonigen, 2010). This holds for almost all skills that we considered. This
implies that life course criminological theories that treat skills as constructs with age-invariant payoffs are missing important variation and should incorporate more dynamic perspectives for skills and personality traits in order to explain age-varying differences in offending behavior.

Second, different skills are important in explaining offending at different stages of the life course. This is in accordance with the economic model of Cunha and Heckman (2007) who suggest that a vector of skills is important in explaining a variety of socioeconomic adulthood outcomes. In psychology the notion that personality is a multidimensional construct is well established (Watson, Clark, & Harkness, 1995). We found for males cognitive skills to be important after the emerging adulthood and social skills important during the adolescent period. For females various social skills are important at different stages in life.

Third, different skills are important for males and females. In addition to differ-
CHAPTER 2: EXPLAINING THE AGE-CRIME CURVE

ences in significance for the payoffs we also found different age-varying payoff patterns. In particular, the patterns for the payoffs from cognitive skills are completely reversed for males and females. For males the payoff decreases with age, whilst for females it increases with age. A similar reversed pattern is found for extroversion. For males extroversion decreases offending during adolescence, while for females it increases offending.

Fourth, the payoffs from the skills are able to explain the aggregate age-crime curve within the populations considered. The model residuals show that differences in childhood skills can almost exclusively explain the pattern of the average age-crime curve. This does not imply that life course transitions from employment, marriage and detention have no influence on individual offending trajectories. It merely implies that mean-level payoffs from differences in underlying skills explain aggregate variation in offending over the life-span.

Based on these four conclusions we can reflect on the self-control theory of Gottfredson and Hirschi (1990) and the age-graded theory of social control (Laub & Sampson, 2003). First, in self-control theory the age-varying explanatory power of skills is explicitly ruled out in the inexplicability hypothesis stated in Hirschi and Gottfredson (1983). There it is argued that age itself is the main explanatory variable and that no underlying combination of sociological and psychological variables can explain the age-crime curve. The results in Figures 2.2 and 2.3 show that for the samples of disadvantaged youths there exists interaction between age and skills, thus providing evidence against the inexplicability hypothesis. Further, in their non-interactive hypothesis they suggest that the same variables are responsible for explaining the cross-sectional differences in offending at any age. In their general theory of crime, or self-control theory, they argue that the responsible variables can be characterized as self-control. As the Figures 2.2 and 2.3 also show that different skills are important at different ages they also provide evidence against the non-interactive hypothesis.

The age-graded social control theory of Laub and Sampson (2003) states that the impact of different social bonds varies over the life span. When translated to a statistical perspective the age-graded social control theory posits a dynamic factor structure for social bonds by arguing that different social bonds can be important at different ages in determining the propensity towards crime. A minor extension would incorporate a similar age- graded perspective for childhood skills. Thus, adopting a dynamic
factor structure for the skills as proposed in this chapter.
Chapter 3

Crime, Employment and Social Welfare: an individual-level study on disadvantaged males

3.1 Introduction

A variety of economic and sociological theoretical mechanisms, such as financial gains, the reduction of inequality and the creation of social bonds, suggest that employment has the potential to reduce criminal behavior (Lageson & Uggen, 2013). Considerable empirical evidence has been found for a negative causal effect of employment on crime (Mustard, 2010). The majority of compelling evidence has been documented by population studies, where labor market prospects and employment stand out as important determinants of crime rates (e.g., Raphael & Winter-Ebmer, 2001; Gould, Weinberg, & Mustard, 2002; Machin & Meghir, 2004; Lin, 2008).

The prevailing question is whether and why the economic or sociological perspective has more weight in explaining the negative relationship between employment and crime. The economic, or rational choice, perspective is documented in Becker (1968), whereas the sociological perspective, which stresses concepts such as social bonds and identity transformations, is discussed in Laub and Sampson (2003). Distinguishing between these perspectives is of vital importance for designing job market programs for offenders, i.e. from a policy perspective. In addition, the theoretical mechanisms
can be used as arguments for the design of welfare policies, that ultimately aim to facilitate reentry into the labor market (Foley, 2011).

Most western societies have implemented a mixture of social welfare policies, such as social insurance and public assistance policies, for those who are not able to find employment and for those who are not capable of working. The effect of such social welfare policies on criminal behavior is not well understood. From an economic perspective welfare payments provide financial gains to the recipients and should therefore reduce the relative returns from criminal behavior (e.g., Becker, 1968; Ehrlich, 1973). However, from a sociological point of view, welfare payments do not stimulate the structure, maturity, responsibility, social bonds and changes in identity that employment provides (e.g., Goodman, 1956; Maruna, 2001; Laub & Sampson, 2003). By including employment, social welfare and crime in a single framework we simultaneously investigate the economic and sociological perspectives on the association between crime and employment.

We investigate the relationships between crime, employment and social welfare benefits for a sample of disadvantaged males, who were institutionalized in a juvenile treatment center in the 1990’s in The Netherlands (e.g., van der Geest et al., 2009, 2011). These individuals show high crime rates, have unstable employment careers and are relatively often the recipient of welfare payments (van der Geest et al., 2011). The sample is not representative for the general population, but provides a unique opportunity to study individual-level interactions between offending, employment and welfare benefits for disadvantaged individuals. The potential monetary returns from crime reducing policies for high-risk individuals are large (Cohen, 1998).

An individual-level empirical analysis of crime, employment and social welfare is complicated as the outcomes are endogenous, which makes standard regression analysis inconsistent. Two main reasons for the endogenous relationship among crime, employment and social welfare are simultaneity and omitted variables.\(^1\) (Davidson &

\(^1\)The omitted variables problem, or selection problem, arises when variables are omitted from the offending-employment/welfare regression that are not randomly correlated with the outcome variable. The corresponding OLS regression parameter estimates become biased whenever this occurs, see Davidson and MacKinnon (2004, Chapter 8). The simultaneity problem, or reverse causality problem as discussed by Ehrlich (1973), arises when offending, employment and welfare outcomes have mutual causal effects on each other. The corresponding OLS regression parameter estimates, resulting from a one-way regression of employment and welfare outcomes on offending outcomes, become biased whenever this occurs.
CHAPTER 3: CRIME, EMPLOYMENT AND SOCIAL WELFARE

MacKinnon, 2004, Chapter 8). Identification strategies that are based on instrumental variables, which are adopted in population studies (e.g., Raphael & Winter-Ebmer, 2001; Gould et al., 2002), are generally not feasible due to a lack of valid instruments for individual-level studies. To avoid problems from endogeneity we simultaneously model the crime, employment and welfare variables. In particular, we propose a dynamic discrete choice model that is able to separate the structural causal effects from the spurious effects (e.g., Heckman, 1981a, 1981c). Implicitly, the model allows for reciprocal effects from crime, employment and social welfare, which capture structural effects and state dependence (e.g., Thornberry & Christensen, 1984; Sampson & Laub, 1993; Nagin & Paternoster, 2000). Ultimately, when considering the model for different types of crime (property and violent), employment (regular and temporary) and social welfare (insurance and assistance), the model enables us to differentiate between the economic and sociological perspectives that aim to explain the relationship between employment and crime.

The remainder of this chapter is organized as follows. We continue this introduction by explaining the basics of the welfare system in The Netherlands. We emphasize that the system shares many properties with other Western societies. In Section 3.2 we discuss the theoretical mechanisms that link crime, employment and social welfare. Here we also explain our identification strategy, which is further formalized in the econometric model that is presented in Section 3.3. The data origins and constructed variables are discussed in Section 3.4. In Sections 3.5 and 3.6 we present the estimation results, while Section 3.7 provides a general discussion for the results.

3.1.1 A primer on the Dutch welfare state

We briefly discuss some aspects of the Dutch welfare state. For a more complete overview we refer to van Oorschot (2006) and de Mooij (2006). The welfare state in The Netherlands has continuously evolved ever since the Second World War. Many different income support policies have been implemented and their peculiarities are periodically reevaluated and adjusted. Together they form an extensive and, compared to other countries, relatively generous social redistribution system, which is under some stress due to an aging population and increased healthcare expenditures (de Mooij, 2006).

The welfare system is designed to reduce inequality by redistributing income in
order to avoid poverty and social exclusion. To achieve this goal a variety of social policies have been implemented. First, minimum wages are compulsory for employers and they are sufficiently high to ensure that the returns from employment are considerable. For example, in 2013 the minimum wage was €726 per month for an 18 year old and €1596 for individuals above 22. The reason for the steep increase in minimum wages between 18 and 23 is that education is encouraged for young adults and high minimum wages for an 18 year old would reduce the relative returns from education.

Second, social welfare policies have been set up for those who are unable to find employment and those who are incapable of working. Two broad categories of social welfare policies can be distinguished, namely insurance policies and public assistance. The category of insurance policies can be split into unemployment insurance and disability insurance. Unemployment insurance consists of payments for those who have lost their job. To qualify one must have been employed for at least 26 weeks out of the 36 weeks prior to losing one job. Also, the performance on the job must not be the reason for the job loss. Conditional on meeting these requirements the individual receives up to 75% of his previous wage for a number of months depending on the number of months that he was previously employed. While receiving unemployment insurance payments the individual is required to regularly apply for jobs and be permanently available for open positions.

Disability insurance is partially covered by employers and partially by the government. Mental or physical illness must always be established by a doctor. For full-time employees the employer is required to pay at least 70% of the wages for a period of 2 years. For those without full time employment the government covers income loss. The amount is again conditional on the previous wage and the number of days worked. Also, there is a maximum of 2 years for this insurance policy, where after one year a complete check up by an independent doctor determines whether the payments will continue. We emphasize that disability insurance is granted for a large number of different illnesses. The vast majority of these are unlikely to make it impossible for individuals to participate in criminal behavior.

Insurance policies are temporary in nature. When individuals do not manage to find employment within the designated period, or remain unable to work, public assistance is available to replace income. Public assistance, the most important being welfare assistance, is meant to assure recipients a minimum income needed for subsistence.
Such benefits do not require proof of anything other than financial need, nor are they conditional on prior employment. Public assistance is paid from general funds and every citizen of the Netherlands is in principle eligible to receive it, unless they are living with a partner or family who provides them with means of subsistence.

By the end of 2013 1.4 million individuals (out of a labor force of 11 million) were receiving some form of social welfare payments. The welfare policies that we discussed above contain the majority of the 1.4 million recipients: unemployment insurance (437,700), disability insurance (97,500) and public assistance (817,900) (CBS, 2012). It is important to note that social welfare payments in The Netherlands are never declined because of criminal history. However, when an individual is incarcerated the eligibility to social welfare stops temporarily.

### 3.2 Crime, employment and social welfare

A vast amount of theoretical mechanisms postulate linkages between employment and crime, see Chalfin and Raphael (2011) and Lageson and Uggen (2013) for recent reviews from economics and sociology. Classical dynamic theories, which are typically used to explain the linkages, include rational choice theory (Becker, 1968), strain theory (e.g., Merton, 1938; Agnew, 1992), social control theory (e.g., Hirschi, 1969; Sampson & Laub, 1993) and labeling theory (e.g., Tannenbaum, 1938; Lemert, 1967). Key elements in these theories include; economic motivations, inequality perceptions, social bonds and reciprocal effects from offending. The relationship between social welfare benefits and criminal behavior is less well understood as the mechanisms depend on the specific type of welfare payment and the eligibility criteria. We show that the relationship between criminal behavior and social welfare in The Netherlands can be largely understood within the classical theories. We discuss the implications of the different theoretical mechanisms for the relationships between crime, employment and social welfare.

Before we discuss the dynamic theoretical perspectives in detail, we briefly mention the self-control theory of Gottfredson and Hirschi (1990). Self-control theory hypothesizes that the correlation between crime, employment and social welfare can be entirely

---

2The other category that is included in the 1.4 million is maternity leave (42,800), which is not included in the current study.
explained by individual characteristics. In particular, it suggests that underlying personal characteristics that select a person into anti-social behavior and delinquency, also select a person into disadvantaged labor market positions, such as welfare and unemployment. In other words, the outcomes are related, but only via preferences and ability, and not via some structural mechanism. A slightly more dynamic perspective for this theory is obtained when assuming that the effect from the preferences and abilities, or the preferences and abilities themselves, can vary with age. For example, Steinberg et al. (2007) show that impulsive behavior, which is typically associated with anti-social behavior, reduces with age.

Dynamic theories imply that there are structural relationships between crime, employment, and social welfare. Within each outcome, structural effects are referred to as the presence of state dependence (e.g., Heckman, 1981a; Nagin & Paternoster, 2000). For example, past experiences from offending may influence contemporaneous outcomes as they can reduce constraints and strengthen incentives to crime (Nagin & Paternoster, 2000). Similar arguments are made for the presence of state dependence in employment in Heckman (1981a). It is important to distinguish these within-outcome dynamic structural effects from the spurious effects discussed above and the cross-outcome structural effects, such as the effects of employment and social welfare on crime, which we discuss next.

First, we consider rational choice theory, or economic theory, which has found widespread support since the seminal contributions of Becker (1968), Ehrlich (1973) and Block and Heineke (1975). Within this framework, criminal behavior is viewed as illegal employment and criminal behavior results from the risk-return trade off between legal and illegal employment (e.g., Ehrlich, 1973; Grogger, 1998). The goal for the individual is to maximize utility, which is typically defined in monetary terms. Some individuals choose crime rather than legitimate employment because they expect to gain more from crime, while taking into account the expected probability and severity of punishment. This implies a trade off between employment and crime, where in each time period the expected returns from the time invested in employment and offending are compared. Low-income and unemployed individuals are hypothesized to have higher offending probabilities as their relative returns from offending are higher. This holds in particular for crimes that lead to financial gains, which make up the majority of total crimes committed (approximately 60% in The Netherlands (CBS,
When viewing welfare benefits solely from a financial perspective, it can be argued within economic theory that welfare benefits should reduce the probability of offending, as they reduce the relative returns from offending. However, the identification of this effect may depend on the consumption patterns of the individuals (Shapiro, 2005). For example, if welfare benefits are spent directly when received it is possible that individuals will supplement their income in the same month via criminal behavior (Foley, 2011). This could make the effect of welfare on crime difficult to identify on a monthly time scale. Also, there is some evidence that the consumption of drugs and alcohol, which is typically associated with crime, increases when welfare checks arrive (Dobkin & Puller, 2007). Overall, economic theory suggests that for the majority of cases welfare benefits will reduce crime as they provide a financial basis for living, but identification depends on the consumption patterns of the individuals.

Second, the economic strain theory proposed by Merton (1938) hinges upon the notion of unfulfilled expectations. In modern western societies, such as The Netherlands, socioeconomic success is a prominent goal that is shared by the majority of the population. At the same time a large proportion of society has little or no prospects for achieving socioeconomic success (Wilkinson & Pickett, 2009). The economic strain theory is generalized in Agnew (1992) to account for more sources, besides economic, for strain. Prominent factors include strain from relationships, housing situations and general psychological well-being. Within this framework, offending, or more general deviant behavior, is seen as a way of expressing frustrations from not living up, or not being able to live up, to expectations. More specifically, for the disadvantaged individuals in society the risk of committing a crime is increased due to a lack of relevant or necessary (material and/or cultural) resources and because of psychosocial reasons, such as higher stress levels or lower life satisfaction.

Strain theory is thus more associated with an individual-level inequality perception, rather than the absolute levels of economic success that are important in the economic theory of crime (Kelly, 2000). Economic and more general social strain theories predict that employment reduces offending, as employment is the standard way in society to obtain an income. This mechanism essentially applies to the majority of crime types, both property related crimes and violent crimes.

According to economic strain theory, social welfare benefits should reduce property
offending as they reduce financial inequality. However, those who receive benefits may still experience feelings of relative deprivation, because their income level is lower when compared to income from employment. Also, other types of crime, such as violent crimes, may be less affected by social welfare payments as these do not remove the frustration of not being able to meet socioeconomic goals by oneself. The responsibility of providing for oneself is in fact taken over by the welfare check and tells the individual that he is not a contributing member of society (Phelps, 1994). Thus, when viewing strain in a broader perspective than economic, strain theory predicts that welfare payments have limited effect on offending. Both economic theory and strain theory are motivation-based in the sense that they seek to explain the structural relationships between crime, employment and welfare from personal motivations.

Third, in contrast, social control theory states that no special motivation is required for criminal behavior (Hirschi, 1969). Everybody is capable of offending, but the majority of individuals are restrained by the social bonds that tie them to society. Delinquency only becomes an option when these social ties are broken or weakened. Sampson and Laub (1993) build on social control theory by applying it to events that occur over the individual’s life. According to their age-graded theory of social control, social bonds lead to informal social control that prevents criminal behavior, but the origins of the bonds may fluctuate with age. During adulthood important social bonds are created by aspects such as employment and intimate relationships. Employment creates structure and allows for supervision by employers and socialization by co-workers (Hirschi, 1969). In addition, employment may increase one’s patience and risk aversion (Becker & Mulligan, 1997), which leads to a lower offending probability. On the other hand, unemployment takes away structure and routines, which may increase the offending probability.

Thus, the crime reducing effect of employment depends on the structure that employment provides for an individual. Also, new relationships that foster social support causing direct or indirect supervision and control are important determinants for the ability of employment to reduce offending (Sampson & Laub, 2005). Employment that does not generate any of the above benefits is unlikely to reduce offending. As such, mainly the quality and stability of employment prevents individuals from criminal behavior. Welfare payments do not generally provide informal social control and can therefore not be expected to contribute to reducing the offending probability.
Fourth, labeling theory focuses on the consequences of contact with the criminal justice system. It suggests that individuals act according to the label that society attaches to them. Being convicted of a crime leads individuals to identify themselves as being a criminal. Labeling theory argues that official interventions, such as incarceration, can cause individuals to commit more crimes. In addition, societal responses to criminal conviction can increase the individual’s perception of himself as a criminal. Labeling theory regards the behavior of the disadvantaged individuals as emanating from a unique culture of disadvantage, whilst economic, strain and social control theories pose that the disadvantaged individuals are acting and adapting in response to prevailing circumstances and hereby imply a large degree of rationality.

The label, or stigma, from convictions and incarceration also affects the probability of finding employment. A large literature has documented difficulties of obtaining legal employment after criminal behavior (e.g., Pager, 2003, 2007; Pager, Western, & Bonikowski, 2009; Apel & Sweeten, 2010; Raphael, 2011). Both legal consequences and social perceptions can lead to difficulties in finding legal employment. The probability of receiving welfare benefits is not affected via labeling theory, as welfare benefits are not granted conditional on resisting from criminal behavior (see Section 3.1.1). If anything, the identity formation caused by criminal behavior will lead to individuals being less reluctant to apply for, or accept, social welfare.

### 3.2.1 Towards identification in a statistical model

Before we discuss our statistical model in detail we discuss our identification strategy that serves to distinguish between the economic and sociological perspectives for the relationship between crime and employment. First, we acknowledge that no single model can distinguish between all the different theoretical mechanism that link crime, employment and social welfare. Given that we observe outcome variables for crime, employment and welfare, a dynamic structural modeling approach is adopted that enables us to find evidence in favor of the different mechanisms. We start by formulating a baseline dynamic model which allows for the simultaneous modeling of both crime, employment and welfare. The simultaneous analysis of all three outcome variables is necessary since we cannot a priori exclude any linkages.

We start by separating structural effects from spurious effects (e.g., Heckman,
Spurious effects are in this sense defined as effects stemming from variables, other than the outcome variables, that create correlation between the dependent variables. We capture these effects by including observed control variables and unobserved statistical control variables. The latter capture the unobserved preferences and abilities, which we allow to vary with age. In this manner we are able to distinguish between the self-control theory of Gottfredson and Hirschi (1990) and the more dynamic structural theories, such as economic, strain, social control and labeling theories. Self control theory prescribes that the spurious effects should explain the entire correlation structure between the outcome variables.

The model contains equations for crime, employment and welfare in which we allow for feedback effects from criminal behavior towards the employment and welfare equations. This allows us to separate the structural effects from employment and welfare on crime from the reciprocal effects. Strong reciprocal effects imply feedback from criminal behavior, either via the criminal justice system or society, and allow us to assess the importance of stigmatization suggested mainly by labeling theory. We acknowledge that disadvantaged individuals typically start their employment careers from an already disadvantaged position, such that initial conditions in our sample will account for a part of the stigmatization (Heckman, 1981b).

Distinguishing between economic, strain and social control theories is more complicated since all three imply that employment should reduce offending. Economic theory and economic strain theory suggest that social welfare should reduce offending as it decreases the relative returns from offending and reduces financial inequality. Social control theory and more general strain theory expect no effect from social welfare on offending. To strengthen our arguments, we use different types of crimes, employment and welfare benefits. By exploiting differences between property and violent offenses, regular and temporary employment and insurance policies and public assistance, we make additional arguments for the economic and sociological perspectives.

3.3 Statistical model

Endogeneity is the key difficulty when aiming to quantify structural relationships among crime, employment and social welfare. Omitted variables and simultaneity are the main factors that impose the endogenous relation among the dependent variables.
CHAPTER 3: CRIME, EMPLOYMENT AND SOCIAL WELFARE

(Davidson & MacKinnon, 2004, Chapter 8). For example, unobserved preferences for employment are likely to be correlated with unobserved preferences for criminal activity. Further, via labeling and stigmatization criminal activity is also likely to impact employment and welfare choices. To address these endogeneity issues we propose a structural dynamic logistic model for separating structural and spurious effects.

Suppose that we observe variables for $N$ individuals and $T$ time periods, where the time periods refer to the monthly age cohorts for the disadvantaged males. We index the individuals by $i$ and the time periods by $t$. Let the dependent variables be denoted by

$$
C_{i,t} = \begin{cases} 
1 & \text{if } \pi_{C,i,t} > 0 \\
0 & \text{if } \pi_{C,i,t} \leq 0
\end{cases},
$$

$$
E_{i,t} = \begin{cases} 
1 & \text{if } \pi_{E,i,t} > 0 \\
0 & \text{if } \pi_{E,i,t} \leq 0
\end{cases},
$$

$$
W_{i,t} = \begin{cases} 
1 & \text{if } \pi_{W,i,t} > 0 \\
0 & \text{if } \pi_{W,i,t} \leq 0
\end{cases},
$$

(3.1)

where $C_{i,t} = 1$ when individual $i$ commits an offense in time period $t$ and $C_{i,t} = 0$ otherwise. Similarly we define $E_{i,t} = 1$ when an individual is employed and $W_{i,t} = 1$ when an individual is receiving welfare payments. A detailed discussion on the construction of these variables is given below, where we also distinguish between different types of offending, employment and social welfare. The vector of dependent variables is given by $Y_{i,t} = (C_{i,t}, E_{i,t}, W_{i,t})'$, for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The net utilities for the offending, employment and welfare outcomes are summarized in $\pi_{i,t} = (\pi_{C,i,t}, \pi_{E,i,t}, \pi_{W,i,t})$, where $\pi_{C,i,t}$ is the net utility that is derived from offending, $\pi_{E,i,t}$ is the net employment utility and $\pi_{W,i,t}$ is the net welfare utility. A similar framework for individual-level decision problems in crime is considered in Durlauf, Navarro, and Rivers (2010).

We model the transformed latent utilities $\theta_{i,t} = \exp \pi_{i,t} / (1 + \exp \pi_{i,t})$ simultaneously by

$$
\theta_{i,t} = \Gamma Y_{i,t-1} + X_{i,t}\beta + \chi_{i,t},
$$

(3.2)

where $\Gamma$ is the $3 \times 3$ matrix that captures the structural effects from offending, employment and welfare from the previous time period, $X_{i,t}$ is a matrix of observed
explanatory variables that is measured by $\beta$ and $\chi_{i,t}$ is the $3 \times 1$ vector of unobserved random effects.

The model aims to separate the structural effects $\Gamma Y_{i,t-1}$ from the observed and unobserved spurious effects that are captured by $X_{i,t} \beta$ and $\chi_{i,t}$. The unobserved random effects are modeled by a factor structure (e.g., Pesaran, 2006; Bai & Ng, 2009; Ahn, Lee, & Schmidt, 2013). For each element $\chi_{j,i,t}$, for $j = C, E, W$, we assume the following specification

$$
\chi_{j,i,t} = \mu_{j,i} f_{j,t}, \quad \mu_{i} = (\mu_{C,i}, \mu_{E,i}, \mu_{W,i})', \quad \mu_{i} \sim N(\delta, \Sigma_\mu),
$$

(3.3)

where the product $\mu_{j,i} f_{j,t}$ is referred to as a factor structure, with $\mu_{i,j}$ being the individual-specific loading coefficient, which is allowed to vary over time according to the common factor $f_{j,t}$. The loadings are assumed normally distributed with mean $\delta$ and variance matrix $\Sigma_\mu$. The off-diagonal elements of the matrix $\Sigma_\mu$ capture the spurious correlation among the dependent variables. We model the factors using cubic spline functions (e.g., Poirier, 1976; Jungbacker, Koopman, & van der Wel, 2014). This specification allows the factors to vary with age, while retaining a parsimonious model specification. We choose different knots for the splines for each yearly age cohort. Experiments with a finer grid for the knots led to similar results. More details are given in Appendix B.

In order to provide some intuition for the structure of the unobserved components $\chi_{j,i,t}$ suppose that $f_{j,t} = 1$, for all $t = 1, \ldots, T$. The model now reduces to a multivariate random effects logistic model, where $\mu_{j,i}$ captures the unobserved mean differences between the individuals for each dependent variable $j$. For example, $\mu_{C,i}$ captures the individual-specific latent preferences for offending. By considering $f_{j,t}$ to be varying over time, we can allow for the unobserved preferences and abilities to vary over time. It is possible to allow for more than one common factor $f_{j,t}$ per outcome variable. However, in our empirical application we found no evidence for multiple factors. We therefore do not discuss this possibility further.

So far we have implicitly assumed that the random unobserved effect is orthogonal to the explanatory variables $X_{i,t}$ and the initial conditions $Y_{i,0}$. In practice this is often an unrealistic assumption (e.g., Chamberlain, 1980; Wooldridge, 2005). Therefore, we explicitly model the correlation between the unobserved components and the
deterministic terms by specifying
\[
\delta = \delta_0 + \left(T^{-1} \sum_{t=1}^{T} X_{i,t} \right) \delta_1 + \lambda Y_{i,0}, \tag{3.4}
\]
where \( \delta_0 \) is the mean vector, \( \delta_1 \) captures the correlation between the individual-specific effects and the means of the explanatory variables and \( \lambda \) captures the correlation between the individual-specific effects and the initial observations. This specification for the individual-specific effects is considered for linear models in Mundlak (1978) and more recently for nonlinear models in Wooldridge (2005).

The complete multivariate logistic panel data model is summarized by equations (3.1), (3.2), (3.3) and (3.4). The coefficients of the matrix \( \Gamma \) are of main interest in our study. These coefficients do not capture contemporaneous effects among the dependent variables. To identify these effects instrumental variables are required, which are not available to us. However, since we have detailed monthly information for offending, employment and social welfare, the structural effects from the previous time period allow us to study dynamic interactions among the dependent variables. See Keane (2010) for a general discussion regarding this approach.

The model parameters are summarized in the vector \( \psi \), which contains the unrestricted elements of \( \Gamma, \beta, \Sigma, \delta_0, \delta_1 \) and \( \lambda \), as well as the knots of the splines. The parameters are estimated by using the Monte Carlo maximum likelihood methods developed in Jungbacker and Koopman (2007) and Mesters and Koopman (2014). A detailed discussion for the methods is provided in Appendix A at the end of this chapter.

3.4 Data

This section outlines the observed data that we use for estimating the parameters of the discrete choice model in Section 3.3. The sample that we consider consists of \( N = 270 \) males who were discharged from a judicial treatment institution in The Netherlands between January 1989 and June 1996. All individuals received treatment during their stay. We only included individuals who stayed in the institution for more than two months and had a complete treatment file. The individuals are observed from age 18
until 32. However, some individuals enter the observational period after the age of 18 because they were still receiving treatment and some individuals leave the sample earlier due to emigration or death.

In The Netherlands juveniles are sent to treatment institutions for various reasons that include: serious behavioral problems, criminal activity and disrupted family situations. Either of these, or a combination, can make it impossible for an individual to remain at home. The median ages in the institution are between 15 and 18. From the age of 12 treatment can be imposed as a criminal law measure. Before the age of 12 treatment can only be imposed as a civil law measure. Eighty percent of our sample was sent to the treatment institution by a civil law measure. The distinction between criminal and civil law says little about the severity of behavioral problems or whether an individual has a conviction prior to treatment in the institution (Wijkman et al., 2006). While in the institution, behavioral problems are treated and low-level education is provided. We emphasize that we do not investigate the effect of the treatment in the institution.

Based on their early encounters with the justice system and their behavioral problems, the individuals can be regarded as constituting a disadvantaged high-risk sample (van der Geest et al., 2009). Typically, disadvantaged individuals are responsible for a large portion of all offenses committed. For all individuals we have detailed registered information on offending, employment and social welfare, from which we construct our dependent variables on a monthly time scale.

### 3.4.1 Crime

The offending and incarceration data are obtained from convictions registered in the Judicial Documentation (JD) abstracts of The Netherlands Ministry of Justice. These are comparable to rap sheets in the US. The abstracts contain information on every case that is sent to the Public Prosecutor’s Office and the verdict that follows from it. They also contain information on date and type of the offense. The abstracts are available for each individual from age 12 and onwards, 12 being the age of criminal responsibility.

We include all property and violent offenses between ages 18 and 32. The property offending category includes crimes such as embezzlement, theft, forgery and counter-
feiting, breaking and entering, burglary, fraud and dealing in stolen property. In total 63.3% of the individuals committed at least one property offense during the observation period. The violent offending category includes assaults, threats, homicides, sexual offenses, robberies and kidnapping. 48.1% of the individuals committed at least one violent offense during the observation period. We refer to CBS (2010) for a complete overview of all the specific crime types that are included in the property and violent offending categories.

The property offenses are denoted by $C_{i,t}^p = 1$ if individual $i$ committed at least one property offense in month $t$ and zero otherwise. Similarly violent offenses are coded as $C_{i,t}^v = 1$, where the index $i$ is for $i = 1, \ldots, N$, with $N = 270$, and the index $t$ is for $t = 1, \ldots, T$, with $T = 168$, and $t = 1$ corresponds to age 18 month 1. Prior offense information, for $t < 1$, indicates that 91.7% of our sample did commit an offense prior to the age of 18.

### 3.4.2 Employment and social welfare

The employment and welfare data is obtained in 2007 from the Ministry of Social Affairs and Employment (SZW). The information consists of individual-level employment and social welfare histories from 1992 onwards. Before 1998 information was only partially administrated. Some employment sectors were excluded. The magnitude of the missing part is unknown. We check the impact by also considering the sample from age 21 onwards, since most individuals were older than 21 in 1998.

For each employment spell we know the exact start and ending date of the contract. Whether a position was full-time or part-time remains unknown to us as we have no information on the exact amount of hours spent working. Also, the SZW database contains no reliable information on wages. The absence of wage information is a minor issue in our study, since The Netherlands is a country with a relatively high minimum wage (see Section 3.1.1).

We construct three different employment variables, which include spells that pertain to any type of employment, spells that pertain to regular employment and spells that pertain to employment via a temporary job agency. The latter distinction allow us to investigate the effect of job stability, since employment through a temporary job agency in the Netherlands often is seasonal or project based, and generally lasts for short spells.
of a few weeks to a few months providing little long-term prospects compared with regular jobs. We code the employment variable as $E_{i,t} = 1$ if individual $i$ is employed in month $t$ and zero else. Similarly, we define $E_{i,t}^r = 1$ if the employment is regular and $E_{i,t}^t = 1$ if the employment is temporary. In our sample 84.4% of the individuals had at least one employment contract, 74.8% had at least one regular employment contract and 67.4% had at least one temporary employment contract. Thus, we may conclude that the majority of our sample has at least some access to the labor market, despite the high level of prior convictions (91.7%).

Three types of social welfare are recovered from the SZW database: unemployment insurance, disability insurance and public assistance. Initially, we take all together because they all imply financial gains for the recipients. We denote the total social welfare outcome by $W_{i,t} = 1$ if individual $i$ is receiving social welfare in month $t$, and zero else. In our sample 61.8% of the individuals received at least once some form of social welfare. Further, we check the robustness for this choice by considering the model with only unemployment insurance ($W_{i,t}^u = 0, 1$), disability insurance ($W_{i,t}^d = 0, 1$), or public assistance ($W_{i,t}^p = 0, 1$) benefits. The insurance benefits and employment are not mutually exclusive as partial employment can in some cases be combined with insurance benefits. 34.4% of the individuals received unemployment insurance benefits at least once, 13.3% received disability benefits at least once and 35.5% received public assistance at least once.

### 3.4.3 Additional control variables

Although the interactions between offending, employment and welfare are of central interest, it is necessary to control for other factors stressed by economic and sociological theories on crime. Therefore, we include binary and count control variables for marriage, divorce and the number of children. The variables are extracted from the municipal registries. The variables are all dynamic and may lead to changes in offending, employment and welfare outcomes (Sampson and Laub (1993); e.g. for marriage and divorce: Bersani, Laub, and Nieuwbeerta (2009); Stolzenberg and D’Alessio (2007)). By including marriage and the number of children we control for the over the years increasing overlap in the ages of marriage, parenthood and employment (Shanahan, 2000). While our central focus is not on these variables we want to include these vari-
ables as controls to limit their influence on the relationships among crime, employment and social welfare. In our sample 22.2\% of the individuals was married (at least once), 8.1\% got divorced and on average individuals had 0.56 children.

A special situation occurs when an individual is incarcerated. In this situation, which occurs at least once for 47\% of the individuals in our sample, it is not possible to be employed or to receive social welfare benefits. Also, offending is less likely during detention. Incarceration severely constrains the individual from making decisions regarding crime, employment and social welfare for the corresponding time periods. However, incarceration does have consequences for future outcome variables (e.g., Apel & Sweeten, 2010; Raphael, 2011). We incorporate incarceration in our model as follows. For the period spent in detention the dependent variables $Y_{i,t}$ are set to missing. For the periods after detention the detention counter variable $D_{i,t}$ (included in $X_{i,t}$) is increased by the number of months spent in detention. This allows us to separate effects from offending and detention\(^3\).

In a small number of cases employment contracts are not terminated during detention. This typically occurs when the detention spells are short. Also, a small number of offenses is committed while being incarcerated. In these situations we consider the detention outcome leading.

### 3.4.4 Summary statistics

Summary statistics for offending, employment, social welfare and the control variables are given in Table 3.1. The average property offending rate is 0.026 per individual per month. The violent offending rate is lower at 0.001. On average around 31\% of the individuals in the sample was employed in each monthly time period. Very few had employment for the entire observation period. The employment rate is similar compared to other studies where high risk offenders are analyzed (e.g., Grogger, 1995; Levitt & Venkatesh, 2001). The majority of the employment is characterized as regular 23.2 \%, while 8.7\% is obtained via temporary job agencies. On average 10\% of the individual monthly periods are related to some form of social welfare. The majority of this, 7.1\%, is labeled as welfare from public assistance. On average 2.3\% claimed

\(^3\)In our empirical study we experimented with different ways of including the detention variable. No qualitative changes in the structural effects were found for different constructions of the detention variable.
unemployment insurance and 0.6% disability insurance.

The summary statistics for marriage, divorce and children indicate that 7.0% of the per person per age periods involve a marriage, 2.5% a divorce and 0.2 children. The ratio between divorce and marriage is almost 1 to 3 which is very high given the young ages.

In Figure 3.1 we show the average time series for the panels of variables. The property offending rate is declining for the older age cohorts. The rate is quite noisy and large spikes are visible. This is caused by the fact that we consider a monthly sampling period. The rate fluctuates between 6% and 0.1% between age 18 and 32. The violent offending rate is equally noisy but does not show a clear declining pattern over age. The level of violent offending is lower when compared to property offending, but for older age cohorts the difference declines.

The employment rate increases until the age of 25. After this age it remains stable around 40%. Two situations are possible. Either the individuals that have acquired sufficient human capital have found their place in the labor market and the individuals that have not have simply dropped out (or are receiving welfare). Alternatively, there are still transitions taking place, but the average just happens to be stuck around 40% since many individuals pass in and out off employment. A quick initial inspection reveals that transitions are still frequent among the older age cohorts.

Interestingly the stagnation in the employment participation rate coincides with an increase in the social welfare rate. This rate increases rapidly between ages 25 and 30. The rate climbs from 0.05 to 0.25. This indicates that at the end of the sample nearly 25% of the individuals who are still in the sample were receiving some form of welfare benefits. The decomposition of social welfare into its components is different in our sample when compared to the decomposition for the general population. In particular, the proportion receiving public benefits is much larger for these disadvantaged males. The incarceration rate, after an initial increase, fluctuates between 5 and 15 % and appears to be declining slightly. One individual in the sample was incarcerated for the entire observation period and therefore does not contribute to the statistical analysis below. The marriage, divorce and parenthood rates increase as the men age.

The overall correlations between the outcome variables are given in Table 3.2. We find that property offending is negatively correlated with all categories of employment and social welfare. The correlation between property offending and total employment
Table 3.1: Summary statistics for all variables for the full sample of individuals (N = 270) and age groups (T = 168).

Table 3.2: Correlations among the dependent variables for the full sample of individuals (N = 270) and age groups (T = 168).

is -0.074 and the correlation between offending and the total welfare variable is -0.026. Regular employment is more strongly negatively correlated with property offending (-0.075) when compared to temporary employment (-0.011). The correlation between violent offending and total employment is negative (-0.024), but zero when looking at only temporary employment. Social welfare and violent offending are slightly positively correlated. The positive correlation is mainly driven by public assistance, whilst the insurance benefits are not correlated with violent offending.
3.5 Trivariate model results

In this section we use the econometric model of Section 3.3 to investigate whether the correlations shown in Table 3.2 stem from structural relationships or from spurious relationships. Further, the parameter estimates aid to distinguish between economic and sociological theories for employment and crime. We first discuss the parameter estimation results for the total employment and social welfare variables which are described in Section 3.4. In Section 3.6 we discuss the results for the different types of employment and social welfare.

The parameter estimates for the trivariate logistic panel data model, which is given in equations (3.1), (3.2) and (3.3), are given in Table 3.3. The full panel of observations is used to calculate the estimates ($N = 270$ individuals and $T = 168$ time periods). We estimated the model parameters separately for property and violent offending using the Monte Carlo maximum likelihood procedure that is discussed in Appendix A.

The estimates show that both structural and spurious effects are present among
property offending, employment and social welfare. In particular, significant effects are found among both the structural parameters $\Gamma$ and the spurious parameters $\beta$, $\delta$ and $\Sigma_\mu$. Positive state dependence is present for all three outcome variables, which shows that past outcomes increase the probability for a new occurrence. For employment and social welfare the coefficients are large (5.627 and 6.354), which indicates that previous employment and social welfare choices have a large impact on subsequent choices for employment and social welfare. State dependence is also present for property offending but the coefficient is lower (1.280).

All cross-equation structural relationships, with the exception of the relationship between crime and social welfare, are statistically significant at the $\alpha = 0.05$ level. The effect of employment on property offending is negative with coefficient -0.424, while the reciprocal effect of property offending on employment is also negative with coefficient -0.902. This indicates that the reciprocal effect has an important role in determining the structural relationship. Also, it implies that a unidirectional approach for modeling property offending and employment is less appropriate (Alessie, Hochguertel, & van Soest, 2004).

The structural relationships between social welfare and property offending are not significant. For the effect of property offending on social welfare this is to be expected as social welfare is not denied because of previous offending. However, we might have expected social welfare to become common after offending as employment possibilities are reduced. The relationship between employment and social welfare is two-way significant with negative signs. The magnitudes are large which implies that transitions in consecutive periods from employment to welfare and vice versa are unlikely.

The control variables marriage, divorce, children and detention have signs as expected. For property offending the number of children has a significant negative effect, with coefficient -0.564. Marriage reduces property offending whilst divorce and incarceration increase the offending probability, albeit not significantly. Both marriage and divorce significantly increase the probability for employment while incarceration reduces the probability for employment. We emphasize that the coefficient for incarceration indicates the marginal effect for one additional month spent incarcerated. None of the control variables significantly affects the social welfare probability.

The most important spurious coefficients are also given in Table 3.3. These show the importance of allowing for cross-section heterogeneity in the model. For all three
dependent variables the variances are high, indicating large mean differences between the individuals. The covariance between property offending and employment is significant and negative. This means that the total negative correlation between property offending and employment, given the control variables, is explained by both structural relationships and spurious relationships. Interestingly the covariance between employment and social welfare is positive. This means that individuals who have an on average higher probability for employment will also have a higher probability for social welfare. We emphasize that this result is perfectly compatible with the negative structural relationships found. In Figure 3.2 we show the estimated factors which are modeled using cubic spline functions. We have scaled the factors with the means $\delta$ of the unobserved heterogeneous effects $\mu_i$. We find that the trend in property offending is declining with age. The factors for employment and social welfare are increasing with age. Effectively, these factors remove the latent common trends, scaled by $\mu_i$, in property offending, employment and social welfare.

Next, we discuss the results for violent offending. Again we find evidence for both structural and spurious effects. The relationships between employment and social welfare remain approximately the same as when property offending was modeled. Therefore, we only discuss the relationships with respect to violent offending. The structural effects indicate that state dependence is also significant for violent offending but the coefficient is smaller (0.892). The causal effect from employment on violent offending is not significant, but the reciprocal effect is large and significant with value -1.151. This shows that violent offending has larger consequences for future employment careers, when compared to property offending. The structural relationships between social welfare and crime remain insignificant.

The control variables for violent offending have the same sign as for property offending. Only the number of children has no significant effect anymore on the violent offending probability. Thus, parenthood does not reduce violent offending. The spurious effects indicate large heterogeneity for the dependent variables. The spurious covariance between employment and offending is not significant for violent offending. In Figure 3.2 we also show the common factor for violent offending. In contrast to property offending we find no clear age trend for the fluctuations of the spurious effects for violent offending. The slope of the factor for violent offending is nearly flat.

Overall we must conclude that the empirical evidence shows a much weaker rela-
CHAPTER 3: CRIME, EMPLOYMENT AND SOCIAL WELFARE

Table 3.3: Parameters estimates for the trivariate logistic panel data model with time-varying individual-specific effects. We consider the full sample of individuals (N = 270) and age groups (T = 168). The * indicates that the coefficient is significant at the \( \alpha = 0.05 \) level.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter</th>
<th>Property offending</th>
<th>Violent offending</th>
</tr>
</thead>
<tbody>
<tr>
<td>State dependence offending</td>
<td>( \Gamma_{1,1} )</td>
<td>1.280*</td>
<td>0.892*</td>
</tr>
<tr>
<td>Employment on offending</td>
<td>( \Gamma_{1,2} )</td>
<td>-0.424*</td>
<td>-0.040</td>
</tr>
<tr>
<td>Social welfare on offending</td>
<td>( \Gamma_{1,3} )</td>
<td>-0.151</td>
<td>0.097</td>
</tr>
<tr>
<td>State dependence employment</td>
<td>( \Gamma_{2,1} )</td>
<td>5.627*</td>
<td>5.630*</td>
</tr>
<tr>
<td>Offending on employment</td>
<td>( \Gamma_{2,2} )</td>
<td>-0.902*</td>
<td>-1.151*</td>
</tr>
<tr>
<td>Social welfare on employment</td>
<td>( \Gamma_{2,3} )</td>
<td>-1.852*</td>
<td>-1.843*</td>
</tr>
<tr>
<td>State dependence social welfare</td>
<td>( \Gamma_{3,1} )</td>
<td>6.354*</td>
<td>6.383*</td>
</tr>
<tr>
<td>Offending on social welfare</td>
<td>( \Gamma_{3,2} )</td>
<td>0.162</td>
<td>0.664</td>
</tr>
<tr>
<td>Employment on social welfare</td>
<td>( \Gamma_{3,3} )</td>
<td>-1.416*</td>
<td>-1.390*</td>
</tr>
<tr>
<td>Marriage on offending</td>
<td>( \beta_{1,1} )</td>
<td>0.214</td>
<td>0.585</td>
</tr>
<tr>
<td>Divorce on offending</td>
<td>( \beta_{1,2} )</td>
<td>0.302</td>
<td>-0.548</td>
</tr>
<tr>
<td>Parenthood on offending</td>
<td>( \beta_{1,3} )</td>
<td>-0.564*</td>
<td>0.287</td>
</tr>
<tr>
<td>Incarceration on offending</td>
<td>( \beta_{1,4} )</td>
<td>0.004</td>
<td>0.001</td>
</tr>
<tr>
<td>Marriage on employment</td>
<td>( \beta_{2,1} )</td>
<td>0.280*</td>
<td>0.291</td>
</tr>
<tr>
<td>Divorce on employment</td>
<td>( \beta_{2,2} )</td>
<td>0.439*</td>
<td>0.459</td>
</tr>
<tr>
<td>Parenthood on employment</td>
<td>( \beta_{2,3} )</td>
<td>-0.061</td>
<td>-0.050</td>
</tr>
<tr>
<td>Incarceration on employment</td>
<td>( \beta_{2,4} )</td>
<td>-0.008*</td>
<td>-0.010*</td>
</tr>
<tr>
<td>Marriage on social welfare</td>
<td>( \beta_{3,1} )</td>
<td>-0.260</td>
<td>-0.260</td>
</tr>
<tr>
<td>Divorce on social welfare</td>
<td>( \beta_{3,2} )</td>
<td>0.369</td>
<td>0.337</td>
</tr>
<tr>
<td>Parenthood on social welfare</td>
<td>( \beta_{3,3} )</td>
<td>-0.201</td>
<td>-0.195</td>
</tr>
<tr>
<td>Incarceration on social welfare</td>
<td>( \beta_{3,4} )</td>
<td>-0.001</td>
<td>-0.006</td>
</tr>
<tr>
<td>Mean offending</td>
<td>( \delta_{0,1} )</td>
<td>-3.829*</td>
<td>-5.214*</td>
</tr>
<tr>
<td>Mean employment</td>
<td>( \delta_{0,2} )</td>
<td>-5.389*</td>
<td>-5.397*</td>
</tr>
<tr>
<td>Mean social welfare</td>
<td>( \delta_{0,3} )</td>
<td>-7.729*</td>
<td>-7.584*</td>
</tr>
<tr>
<td>Variance offending</td>
<td>( \Sigma_{v,1,1} )</td>
<td>1.787*</td>
<td>1.857*</td>
</tr>
<tr>
<td>Covariance offending employment</td>
<td>( \Sigma_{v,2,1} )</td>
<td>-0.462*</td>
<td>-0.299</td>
</tr>
<tr>
<td>Covariance offending social welfare</td>
<td>( \Sigma_{v,3,1} )</td>
<td>-0.492*</td>
<td>0.369</td>
</tr>
<tr>
<td>Variance employment</td>
<td>( \Sigma_{v,2,2} )</td>
<td>2.531*</td>
<td>2.549*</td>
</tr>
<tr>
<td>Covariance employment social welfare</td>
<td>( \Sigma_{v,3,2} )</td>
<td>1.184*</td>
<td>1.012*</td>
</tr>
<tr>
<td>Variance social welfare</td>
<td>( \Sigma_{v,3,3} )</td>
<td>3.835*</td>
<td>3.480*</td>
</tr>
</tbody>
</table>

The relationship between employment and violent crime. Only the structural effect of violent offending on future employment is significant. Multiple reasons exist for such a weak relation, but an important factor is that our sample is relatively small and violent offenses are rare, even for the high risk sample under consideration.

3.5.1 Robustness

In order to highlight some of the features of the statistical model we briefly discuss some additional results from restricted model specifications. In particular, in Table 3.4 we show the structural parameters that are obtained by (a) restricting the model to be uni-directional and (b) by considering the model with time-invariant individual-specific effects. The former reduces the model to the standard random effects model.
Figure 3.2: Smooth factors for offending, employment and social welfare

with lagged explanatory variables (Baltagi, 2005), while the latter considers equation (3.3) for $f_{j,t} = 1$ for all $j$ and $t$. We omitted the parameters for the control variables from Table 3.4.

The uni-directional models tend to overstate the cross-outcome structural effects. For property and violent offending the effects of employment and social welfare are larger in magnitude, but the statistical significance remains the same. The uni-directional model for employment shows that this is not always the case. The effects of property and violent offending on employment are not significant anymore, which indicates that bi-directional models for the effect of employment on offending can reveal additional structural relationships.

When the individual-specific effects are held constant with age we find that the structural effects between employment and property offending become larger in magnitude. Also, the structural effects between social welfare and property offending are now significant. This indicates that the correlation that was previously attributed to the spurious time-varying effect is now picked up by the social welfare variable. This
clearly illustrates the importance of allowing the latent spurious effects to vary with age. For violent offending the changes are smaller. This is not surprising since the slope of the unobserved trend in violent offending is lower in magnitude (see Figure 3.2).

Additional robustness tests that we performed included: implementing the detention variable as a regular control variable and estimating the model for time periods that corresponded to ages 21 and older. None of these gave qualitatively different results when compared to our preferred model specification that is discussed above. We therefore do not show these results.

### 3.6 Different types of employment and social welfare

Next, we discuss the results for the different types of employment and social welfare. In particular, we re-estimate the parameters of the trivariate structural model of Section 3.3 for regular and temporary employment as well as unemployment insurance,
disability insurance and public assistance. When we change a particular dependent variable we keep the other dependent variables similar as in Section 3.5. This allows us to compare the results for the structural parameters to those given in Table 3.3.

The parameter estimates for the regular and temporary employment variables are given in Table 3.5. We only show the structural parameter estimates, since these are of main interest. The effect of regular employment on property offending is significant and almost twice as large when compared to the results for total employment in Table 3.3 (-0.736 vs -0.424). Recall that regular employment refers to employment that is registered on the payroll of the employer. The effect of temporary employment (via a temporary job agency) on property offending is not significant. The effects of both regular and temporary employment on violent offending are insignificant. The state dependence parameters for employment indicate that regular employment is also more persistent when compared to temporary employment (6.282 vs 5.257).

The reciprocal effects of employment on offending show an interesting separation. Both property and violent offending have no significant effect on regular employment, while the effects are large and significant for temporary employment. A tentative conclusion for this finding is that employees with a regular employment contract can be much harder to fire when compared to employees with a temporary employment contract. Further, employers might be more reluctant to fire permanent employees as it involves additional costs for hiring and training new employees.

The parameter estimates for the different social welfare categories are given in Table 3.6. For the insurance policies, unemployment and disability, we find no differences when compared to the total welfare category. The structural relationships between the insurance polices and property and violent offending are not significant. For the public assistance category we find two interesting structural relationships. First, public assistance significantly lowers the probability of property offending. The coefficient is large in magnitude (-0.490). This is comparable to the coefficient for employment (-0.448), but the 95% confidence bounds are larger ([-0.956,-0.024] for public assistance vs [-0.687,-0.209] for employment). Second, violent offending significantly increases the probability for public assistance (1.189). This indicates that individuals who commit violent offenses are more likely to receive public assistance benefits.
CHAPTER 3: CRIME, EMPLOYMENT AND SOCIAL WELFARE

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Regular employment</th>
<th>Violent offending</th>
</tr>
</thead>
<tbody>
<tr>
<td>State dependence offending</td>
<td>$\Gamma_{1,1}$</td>
<td>$0.907^*$</td>
</tr>
<tr>
<td>Employment on offending</td>
<td>$\Gamma_{1,2}$</td>
<td>$0.027$</td>
</tr>
<tr>
<td>Social welfare on offending</td>
<td>$\Gamma_{1,3}$</td>
<td>$0.129$</td>
</tr>
<tr>
<td>State dependence employment</td>
<td>$\Gamma_{2,2}$</td>
<td>$-1.027^*$</td>
</tr>
<tr>
<td>Offending on employment</td>
<td>$\Gamma_{2,1}$</td>
<td>$1.157^*$</td>
</tr>
<tr>
<td>Social welfare on employment</td>
<td>$\Gamma_{2,3}$</td>
<td>$0.148$</td>
</tr>
<tr>
<td>State dependence social welfare</td>
<td>$\Gamma_{3,3}$</td>
<td>$6.675^*$</td>
</tr>
<tr>
<td>Offending on social welfare</td>
<td>$\Gamma_{3,1}$</td>
<td>$0.252$</td>
</tr>
<tr>
<td>Employment on social welfare</td>
<td>$\Gamma_{3,2}$</td>
<td>$-1.049^*$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Temporary employment</th>
<th>Violent offending</th>
</tr>
</thead>
<tbody>
<tr>
<td>State dependence offending</td>
<td>$\Gamma_{1,1}$</td>
<td>$0.907^*$</td>
</tr>
<tr>
<td>Employment on offending</td>
<td>$\Gamma_{1,2}$</td>
<td>$0.027$</td>
</tr>
<tr>
<td>Social welfare on offending</td>
<td>$\Gamma_{1,3}$</td>
<td>$0.129$</td>
</tr>
<tr>
<td>State dependence employment</td>
<td>$\Gamma_{2,2}$</td>
<td>$-1.027^*$</td>
</tr>
<tr>
<td>Offending on employment</td>
<td>$\Gamma_{2,1}$</td>
<td>$-1.053^*$</td>
</tr>
<tr>
<td>Social welfare on employment</td>
<td>$\Gamma_{2,3}$</td>
<td>$1.157^*$</td>
</tr>
<tr>
<td>State dependence social welfare</td>
<td>$\Gamma_{3,3}$</td>
<td>$6.675^*$</td>
</tr>
<tr>
<td>Offending on social welfare</td>
<td>$\Gamma_{3,1}$</td>
<td>$0.252$</td>
</tr>
<tr>
<td>Employment on social welfare</td>
<td>$\Gamma_{3,2}$</td>
<td>$-1.049^*$</td>
</tr>
</tbody>
</table>

Table 3.5: Structural parameter estimates for the trivariate logistic panel data model for regular and temporary employment. We consider the full sample of individuals ($N = 270$) and age groups ($T = 168$). The * indicates that the coefficient is significant at the $\alpha = 0.05$ level.

3.7 Discussion and conclusion

Several theoretical mechanisms predict a negative structural effect of employment on crime (e.g., Ehrlich, 1973; Merton, 1938; Laub & Sampson, 2003). The objective of this chapter was to distinguish between economic and sociological theories for explaining the relationship between employment and crime. The role of social welfare was used as an identifying mechanism. The economic theories included rational choice theory and economic strain theory, whereas social control theory and general strain theory are sociological theories that we considered. Further, we used labeling theory to predict a negative structural effect of offending on employment.

A distinction between economic and sociological theories was found for the relationship between offending and social welfare. When arguing that social welfare does not provide the social structure and social bonds that are associated with employment, sociological theories imply that social welfare should not reduce offending. In contrast, from an economic point of view the improved financial position that results from receiving welfare payments should reduce the offending probability relative to not receiving an income. This would hold mainly for offenses that provide financial gains.
We used an individual-level modeling approach to separate the theoretical mechanisms in our empirical study. In particular, we argued if social welfare is found to structurally reduce offending this would support the idea that providing the means to pursue economic goals is an important factor in explaining desistance. By contrast, if we find no structural effects from social welfare on offending, this is regarded as evidence in favor of sociological theories which assume that employment provides more than an income. This reasoning holds if employment itself, within the same sample, reduces offending. Even with such a trivial identifying scheme, the separation of structural and spurious effects remains a challenging task. Structural relationships can operate in both directions and spurious effects can vary with age. We proposed a dynamic structural model that separates aspects such as state dependence, reciprocal
CHAPTER 3: CRIME, EMPLOYMENT AND SOCIAL WELFARE

effects and unobserved heterogeneity among variables for offending, employment and social welfare.

Our results show that a part of the correlation between employment and crime is spurious. In particular, the covariance between employment and property offending was significant and negative. This implies that self-control theory (Gottfredson & Hirschi, 1990), which assumes that individual preferences and abilities that select individuals into offending also select individuals into unemployment, may have a role in explaining the correlation.

In addition, we found significant structural effects between employment and crime. This implies that the selection mechanism of self-control theory cannot be the only contributing mechanism to the negative correlation between employment and crime, but that the classical dynamic theories, such as rational choice, strain, social control and labeling, may also play a role. First, large and significant structural effects from offending on employment are in line with the idea that stigmatization and labeling theory are important in explaining the correlation. The effects from offending on employment are larger in magnitude for violent offending when compared to property offending. It is unclear whether these reciprocal effects are due to legal or societal responses to criminal behavior. Experimental and legal studies are more suited to distinguish between these two explanations (Pager et al., 2009).

Second, the structural effect of employment on property offending is large and significant, whereas no significant effects are found for the effect of employment on violent offending. However, the coefficients for the structural effect of employment on violent offending are all negative, which possibly indicates that the sample size is not sufficient to identify the effect. The negative effect of employment on property offending was mainly caused by regular employment. In particular, when separating the employment category into regular and temporary employment we found that only regular employment significantly reduced property offending. One possible explanation is that the stability of employment, which is associated with regular employment, is important in explaining the negative effect. This by itself does not allow us to distinguish between sociological and economic theories as both social bonds and financial returns are likely to increase with regular employment.

Therefore we turn to the structural effects from social welfare. Overall, we found no significant structural relationships between social welfare and crime. This holds
for both property and violent crimes. However, when we separated the social welfare categories into insurance policies and public assistance we found that public assistance significantly lowers the probability for property offending. More importantly, we found that the coefficients are roughly of the same magnitude. Public assistance can be regarded as the safety net for the most disadvantaged individuals, on which individuals can rely if insurance policies are unavailable or exhausted. Our findings show that for these most disadvantaged individuals the financial returns from public assistance do provide protection against property offending. This points to a substitution effect between property offending and public assistance as hypothesized by the economic theories of crime (e.g., Becker, 1968; Ehrlich, 1973; Merton, 1938). Since public assistance does not provide any reasonable social bonds it is hard to argue in favor of the sociological theories for explaining the relationship with respect to property offending.

None of the social welfare categories was able to significantly reduce violent offending. This finding can, similarly as for employment, be attributed to a lack of identification due to the small sample size. However, we cannot rule out the possibility that the sociological perspective, which includes aspects such as social bonds and identity transformations are indeed important in explaining the relationship between violent offending and employment.

Overall our results indicate the importance of self-control, labeling and economic theory in explaining the relationship between employment and crime for disadvantaged individuals. The mixture of empirical evidence highlights the complexity of the theoretical relationship between employment and crime.

### Appendix A: Estimation method

In this appendix we discuss the Monte Carlo maximum likelihood method that is used to estimate the parameters of the logistic trivariate panel data model that is discussed in Section 3.3. The methodology is discussed for a more general panel data model in Mesters and Koopman (2014). The parameters are summarized in the vector $\psi$.

We summarize the vector of dependent variables for individual $i$ in time period $t$ by $Y_{i,t} = (C_{i,t}, E_{i,t}, W_{i,t})'$, which is thus a $3 \times 1$ vector of binary variables. The loglikelihood for the observations $Y = \{Y_{i,t}\}_{i=1,...,N \ t=1,...,T}$ is defined as $\ell(\psi; Y) = \log p(Y; \psi)$, where $p(Y; \psi)$ is the joint density of all observations. In the presence of
the random effects $\mu_i$, defined in (3.3), we can express the joint density as a high dimensional integral as follows

$$ p(Y; \psi) = \int p(Y, \mu; \psi) \, d\mu = \int p(Y|\mu; \psi) p(\mu; \psi) \, d\mu, \tag{3.5} $$

where $\mu = \{\mu_i\}_{i=1}^{N}$ and $p(\mu; \psi)$ is defined in equation (3.3). The conditional density $p(Y|\mu; \psi)$ for the trivariate model can be written as

$$ p(Y|\mu; \psi) = \prod_{i=1}^{N} \prod_{t=1}^{T} p(Y_{i,t}|\mu_i; \psi), $$

where

$$ \log p(Y_{i,t}|\mu_i; \psi) = \sum_{j=C,E,W} Y_{j,i,t} \theta_{j,i,t} - \log(1 + \exp \theta_{j,i,t}), $$

where $\theta_{j,t} = (\theta_{C,i,t}, \theta_{E,i,t}, \theta_{W,i,t})$ is given in (3.2) and $Y_{j,i,t}$ corresponds to the outcome variables $Y_{C,i,t} = C_{i,t}$, $Y_{E,i,t} = E_{i,t}$ and $Y_{W,i,t} = W_{i,t}$ (Durbin & Koopman, 2012, e.g. Section 10.3).

As $p(Y|\mu; \psi)$ corresponds to a logistic binary density no closed form solution exists for the high dimensional integral in (3.5). Instead we follow the conventional literature and solve the integral using Monte Carlo methods. We refer to Cappé et al. (2005) and Durbin and Koopman (2012, Part 2) for general introductions into these methods. A simple Monte Carlo estimate is obtained by drawing $S$ samples of $\mu$ from $p(\mu; \psi)$ and computing the average

$$ \hat{p}(Y; \psi) = S^{-1} \sum_{s=1}^{S} \sum_{i=1}^{N} \sum_{t=1}^{T} p(Y_{i,t}|\mu_i^{(s)}; \psi), $$

where $\mu_i^{(s)}$ denotes the $s$th sample from $p(\mu; \psi)$. From the law of large numbers it follows that $\hat{p}(Y; \psi) \to p(Y; \psi)$ as $S \to \infty$. However, the simple estimate requires many draws $S$ before convergence is achieved. This follows as the density $p(\mu; \psi)$ does not account for the observations $Y$.

More efficiency can be obtained by sampling sequences for $\mu$ from an appropriate importance density (Ripley, 1987). For the construction of an adequate importance density we follow Jungbacker and Koopman (2007) and Mesters and Koopman (2014).
The general importance sampling representation for the trivariate model is given by

\[ p(Y; \psi) = \int_{\mu} \frac{p(Y|\mu; \psi)p(\mu; \psi)}{g(\mu|Y)} g(\mu|Y) \, d\mu, \]

where \( g(\mu|Y) \) is the importance density. When applying Bayes rule to the right hand side we obtain

\[ p(Y; \psi) = g(Y) \int_{\mu} \frac{p(Y|\mu; \psi)}{g(Y|\mu)} g(\mu|Y) \, d\mu, \]

where we have imposed \( g(\mu) = p(\mu) \). A Monte Carlo estimate for the importance sampling representation is given by

\[ \hat{p}(Y; \psi) = g(Y) \sum_{s=1}^{S} \frac{p(Y|\mu^{(s)}; \psi)}{g(Y|\mu^{(s)})}, \]

where samples \( \mu^{(s)} \) are drawn independently from importance density \( g(\mu|Y) \).

We choose \( g(\mu|Y) \) to follow a Gaussian density with mean equal to the mode of \( p(\mu|Y) \) and variance equal to the curvature around the mode. An instrumental basis for \( g(\mu|Y) \) that allows us to obtain the mode is given by

\[ z_i = \mu_i + u_i, \quad u_i \sim NID(0, D_i), \]

where \( z_i \) and \( D_i \) are obtained by the following Gauss-Newton algorithm.

**Algorithm**

1. Initialize \( \mu = \mu^* \);
2. Given \( \mu^* \); compute

   \[ D_i = -\left[ \sum_{t=1}^{T} \frac{\partial^2 \log p(Y_{i,t}|\mu^*_i; \psi)}{\partial \mu^*_i \partial \mu^*_i} \right]^{-1}, \]

   and

   \[ z_i = \mu^*_i + D_i \sum_{t=1}^{T} \frac{\partial \log p(Y_{i,t}|\mu^*_i; \psi)}{\partial \mu^*_i}, \]

   for \( i = 1, \ldots, N \);
CHAPTER 3: CRIME, EMPLOYMENT AND SOCIAL WELFARE

3. Update $\mu^*$ by computing $E_g(\mu | z)$ based on $z_i = \mu_i + u_i$ and $u_i \sim NID(0, D_i)$; 

4. Iterate between (2) and (3) until convergence.

Convergence of the algorithm is typically quick (4-5 iterations). The derivatives in step (2) are given in Durbin and Koopman (2012, Part 2). After convergence we have obtained the mode of $p(\mu | Y; \psi)$ and we can sample $S$ times from the importance density $g(\mu | Y) \equiv g(\mu | z)$, where $g(\mu | z)$ is a Gaussian density where the mean and variance are implied by $z_i = \mu_i + u_i$ and the distribution of $\mu_i$ given in (3.3). Using these samples we construct the Monte Carlo likelihood. The resulting likelihood estimate $\hat{p}(y; \psi)$ is optimized with respect to parameters $\psi$ by numerical methods (Nocedal & Wright, 1999). This is done while using the same random numbers and the same number of draws $S$ in each iteration.

Appendix B: Factor splines

In this appendix we provide the details for the construction of the cubic splines that we use to model the factors. More details for methods using splines can be found in Poirier (1976). In principal, it is possible to treat all the factors $f_{j,t}$ as deterministic parameters and estimate them along with the other parameters. However, since the time series dimension is $T = 168$ this would lead to difficulties in optimizing the likelihood using numerical methods.

To avoid this problem, we make the assumption that the individual preferences and abilities vary smoothly with age. This allows us to fit cubic splines for the factors, which rely on a smaller number of parameters. In particular, we seek a subset of $K$ knots denoted by $\tilde{R}_{R(l),t}$, for $l = 1 \ldots, K$, where $R(l) \in \{1, \ldots, T\}$. The locations $R(l)$ of the knots are increasing with age; i.e. $R(1) < R(2) < \ldots < R(K)$. Between these knots we fit cubic polynomial functions to approximate the factors that lie between the knots. The knots $\tilde{R}_{R(l),t}$ are estimated along with the other parameters. The location of the knots can be determined in a variety of ways (Jungbacker et al., 2014). In this paper we set the locations equal to the first month of every age year. Thus, we take in total 15 knots with are placed at age 18 month 1, age 19 month 1 etc. The final knot is for age 31 month 12.
Chapter 4

Childhood Skills, Signals and Socioeconomic Adulthood Outcomes for Disadvantaged Youths

4.1 Introduction

Suppose that an individual is born in a low socioeconomic position and has obtained limited cognitive and non-cognitive skills during childhood. Further, suppose that childhood has produced early contact with law enforcement agencies and low levels of education. We propose an empirical framework to investigate to what extent such disadvantaged childhood circumstances have a lasting impact on subsequent socioeconomic adult outcomes, and whether life course transitions during adulthood, such as those from intimate relationships and employment, additionally influence adult outcomes.

The decomposition of adult outcomes that we propose aims to separate the influences of the following three established findings. First, a growing literature has shown that childhood skills substantially impact adult life outcomes, see Cunha, Heckman, Lochner, and Masterov (2006), Heckman (2008), Almond and Currie (2011) and Heckman and Kautz (2013) for recent reviews of the literature. Multiple skills, includ-
ing cognitive and non-cognitive skills\(^1\), are deemed important in determining successful adult outcomes. These skills are formed during childhood by a dynamic process that is considered self-productive and complementing in its arguments\(^2\) (Cunha & Heckman, 2007). Second, childhood signals from education and criminal justice system contacts provide additional information based on which employers, potential romantic partners and law enforcement agencies can make decisions (e.g., Weiss, 1995; Card, 1999; Heckman, Lochner, & Todd, 2006; Pager, 2007). In this perspective the childhood signals are functions of the childhood skills and can affect adult outcomes via labeling and stigmatization (e.g., Lemert, 1967). Third, life course and human capital theories suggest that adult outcomes may mutually influence each other during adulthood (Becker, 1993; Laub & Sampson, 2003; Green, 2010). For example, several studies have shown crime reducing effects from investments in marriage and employment (e.g., Ehrlich, 1973; Sampson, Laub, & Wimer, 2006; Lageson & Uggen, 2013).

In this chapter we propose a reduced-form nonlinear dynamic panel data model which decomposes adult outcomes into payoffs from childhood skills, childhood signals and dynamic contagion from adult outcomes. In doing so we assess the importance of (a) childhood factors and (b) adult life transitions, in a dynamic setting. Additionally, the decomposition enables us to study adulthood multiplier effects from investments in childhood skills. The childhood skills include cognitive and non-cognitive skills which are identified using a linear factor model (e.g., Heckman, Stixrud, & Urzúa, 2006). The childhood signals include education levels and information from criminal records. We model the childhood signals as functions of the childhood skills which enables us to distinguish between the effects from the signals over and above the effects from the skills on the adult outcomes. We model the socioeconomic adolescent and adult outcomes between ages 16 and 32 by arbitrary non-Gaussian densities, that are defined conditional on a vector of latent adult signals. Each age-varying adult signal includes the childhood skills, the childhood signals and dynamic payoffs from adulthood investments. Despite the large number of components that we include in the adult signals,

\(^1\)Non-cognitive skills are also referred to as social skills or personality traits, see Borghans et al. (2008). In this paper we follow the majority of the economic literature and use the term non-cognitive skills.

\(^2\)Self-productivity refers to the idea that skills that are acquired in one period can have lasting effects on future skills. Dynamic complementarity suggests that skills from one period can raise the returns from investments in later periods.
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

we acknowledge that unobserved variables are likely to remain present. To capture the spurious persistence in the adult outcomes that is caused by missing information we explicitly model individual-specific serially correlated error terms (Keane, 1994; Heiss, 2008). The parameters of the complete model are estimated by Monte Carlo maximum likelihood methods. In particular, we extend the importance sampling methods of Mesters and Koopman (2014) to allow for augmented factor structures and multiple mixed-measurement observations per individual per age period.

We estimate the parameters of our statistical model using data for samples of disadvantaged males and females who were institutionalized in a juvenile treatment facility in the 1990s in The Netherlands 3 (e.g., van der Geest et al., 2009, 2011; Mesters et al., 2014). These individuals started their adult life from a severely disadvantaged position, cumulatively obtained during childhood as a result of low parental investments and supervision, early contact with law enforcement agencies and low levels of education. The adult outcomes are characterized by high crime rates and drug use, unstable employment careers and high dependence on social welfare (van der Geest et al., 2011; Verbruggen, Blokland, & van der Geest, 2012; Mesters et al., 2014). Despite the on average poor outcomes we observe substantial heterogeneity in adult outcomes, which raises the questions during which stage (childhood or adulthood) these differences have occurred and during which stage investments might best be made to improve the adult life outcomes of such disadvantaged youths.

We combine data from the files from the juvenile treatment facility with retrospective interview data and official register data. The treatment files, which contain a large number of measurements relating to cognitive test scores and personality traits, are used to determine cognitive and non-cognitive skills (e.g., Heckman, Stixrud, & Urzúa, 2006). The retrospective interviews that were conducted between ages 30 to 40 provide information on education levels and adult outcomes. The interview included a life-history calendar which provided age-specific details for the adult outcomes between ages 16 and 32. We supplemented the interview data with official registered employment, social welfare and criminal record data. The combination of these sources allows us to include adult outcome variables for crime, employment, social welfare, drug use and intimate relationships.

3A country which is characterized by a generous social welfare system and relatively low income-inequality (de Mooij, 2006; Wilkinson & Pickett, 2009).
This paper contributes to the literature in four ways. The first and main contribution is that we study the influence of childhood factors on adult outcomes in a dynamic model for the adulthood. This in contrast to the majority of empirical studies where the adult outcomes are modeled static (e.g., Heckman, Stixrud, & Urzúa, 2006; Cunha & Heckman, 2008; Cunha et al., 2010; Lindqvist & Vestman, 2011). The dynamic approach allows for the separation of effects from (a) childhood skills and signals and (b) adult life-course transitions on socioeconomic adult outcomes. Further, this enables us to trace shocks from childhood skills through adulthood while taking adult life transitions into account. We emphasize that both the childhood factors and the dynamic complementing influences from adult outcomes are likely to have some merit. The interest for public policy is in the relative explanatory power of each component and their interaction. For example, suppose that the adult outcomes predominantly explain each other in a dynamic setting and that childhood skills have little influence, interventions in adulthood can then have contagion effects on other life course outcomes, making them a suitable target for policy intervention. On the other hand, if childhood skills and signals persistently explain adult outcomes, potentially large improvements can be achieved by childhood interventions. Further, the interaction between contagion effects from adult outcomes and childhood interventions can create multiplier effects that exceed the marginal payoffs from childhood interventions. Only a dynamic setting for the adult outcomes can decompose these influences.

Second, the childhood factors are separated into childhood skills (cognitive and non-cognitive) and childhood signals (education and criminal record). This allows us to distinguish between the effects from skills and stigmatization on adult outcomes. Third, we consider samples of youths, males and females, which are characterized by their severely disadvantaged childhood. For these individuals we consider multiple adult outcomes of which a number have received limited attention in previous studies. Examples include variables for intimate relationships, social welfare and drug use. We emphasize that the potential monetary returns from policy interventions for disadvantaged youths are large (Cohen, 1998). Fourth, we extend the econometric parameter estimation methodology that is developed in Jungbacker and Koopman (2007) and Mesters and Koopman (2014) to include factor structures and multiple outcomes per individual.

The results from our study can be summarized as follows. The parameter estimates
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

indicate that for males cognitive and non-cognitive skills have persistent effects on adult outcomes for crime and employment. For females non-cognitive skills are more important in explaining adulthood offending and employment. The education signals have additional persistent effects on employment and social welfare, whereas the criminal record signal has lasting effects on adulthood offending and drug use for both males and females. Simulation results show that minor improvements in non-cognitive childhood skills (1 SD positive shock) on average reduce serious offending per age-year between ages 17 and 32 by 4.6% (m) and 2.3% (f) and increase employment by 5.4% (m) and 5.2% (f). Minor improvements of cognitive skills (1 SD positive shock) reduce serious offending by 2.6% (m) and 0.1% (f) and increase employment by 4.1% (m) and -1.0% (f).

A limitation of our study is that we only have measures on the cognitive and non-cognitive skills from one period. That is the period that the individuals were institutionalized. We are thus not able to model the dynamics in the accumulation of childhood skills. When this information would be available the model of Cunha et al. (2010) could be appended to our model. While our research goals do not depend on these dynamics, they would strengthen the validity of the latent cognitive and non-cognitive skill factors. Also, and common to most studies that aim to quantify effects from childhood skills on adult outcomes, our sample sizes are relatively small (N = 116 males and N = 132 females), see also Heckman, Moon, Pinto, Savelyev, and Yavitz (2010). In order to minimize the influence of the small sample sizes we propose a one-step estimation method for the model parameters, thus circumventing efficiency loss from two-step methods, such as those considered in for example Lochner (2004).

The remainder of this chapter is organized as follows. In Section 4.2 we discuss the statistical model that decomposes the adult outcomes into childhood skills, signals and dynamic contagion effects from adult outcomes. In Section 4.3 we discuss the origins of our sample of disadvantaged youths and detail the construction of the variables. Section 4.4 presents the parameter estimation results which are used in Section 4.5 to investigate the dynamic implications of our model. Our conclusions are discussed in Section 4.6.
4.2 The statistical model

The statistical model that we consider can be viewed as a factor-augmented nonlinear dynamic panel data model. In contrast to the factor-augmented vector autoregressive model proposed in Bernanke, Boivin, and Eliasz (2005) we augment the factors in the cross-section dimension instead of the time series dimension and the observations are modeled by non-Gaussian densities. In Sections 4.2.1 and 4.2.2 we summarize the model for the childhood from which we obtain the cognitive and non-cognitive skills as well as the childhood signals. In Section 4.2.3 we discuss the dynamic model for adulthood. This model separates the adulthood outcomes into effects from childhood skills, childhood signals and dynamic contagion in adult outcomes. The identification and estimation of the factors and model parameters is outlined in Section 4.2.4.

4.2.1 Childhood skills

Following the theoretical model in Cunha and Heckman (2007) we assume that during childhood the individuals cumulatively acquire a vector of skills that can be separated into two components that represent cognitive and non-cognitive skills. The inputs in this process are parental investments and initial birth endowments which are possibly determined by parental skills. The accumulation process is based on self-reinforcing mechanisms and the dynamic complementarity of skills. We assume that for individual $i$ the result of this cumulative process at the end of childhood can be summarized in a cognitive skill $f_{i,C}$ and a non-cognitive skill $f_{i,N}$.

We extract the latent skills $f_{i,C}$ and $f_{i,N}$ from a linear factor model based on measurements from the treatment files of the individuals from the juvenile treatment facility. These measurements include separate sets for cognitive and non-cognitive skills which are discussed in detail in Section 4.3. At the time of constructing the measurements the individuals are institutionalized and there is slight variation in age and the pursued level of schooling. Since age and schooling level are important determinants for cognitive and non-cognitive skills, we model the factor model coefficients as functions of age and the level of schooling, see also Heckman, Stixrud, and Urzúa (2006).

Let $Z_{i,C}$ and $Z_{i,N}$ denote the vectors of measurements for individual $i$ on cognitive and non-cognitive skills. The vectors are of length $K_j$, with $K_j \geq 2$, for $j = C,N$. 
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

Following Heckman, Stixrud, and Urzúa (2006) the factor model for the childhood skills is given by

$$Z_{i,C} = \mu_C(E_{i,0}, \tau_{i,0}) + \alpha_C(E_{i,0}, \tau_{i,0})f_{i,C} + \epsilon_{i,C}(E_{i,0}, \tau_{i,0})$$  \hspace{1cm} (4.1)

and

$$Z_{i,N} = \mu_N(E_{i,0}, \tau_{i,0}) + \alpha_N(E_{i,0}, \tau_{i,0})f_{i,N} + \epsilon_{i,N}(E_{i,0}, \tau_{i,0})$$  \hspace{1cm} (4.2)

where $\mu_j$ is the $K_j \times 1$ mean vector, $\alpha_j$ is the $K_j \times 1$ loading vector, $f_{i,j}$ is the factor score and $\epsilon_{i,j}$ is the $K_j \times 1$ mean zero disturbance vector, for $j = C, N$. The model parameters are modeled as linear functions of education level at the time of treatment $E_{i,0}$ and age at the time of treatment $\tau_{i,0}$. For example, for individual $i$ the mean is modeled as $\mu_C(E_{i,0}, \tau_{i,0}) = \kappa_0 + \kappa_1 E_{i,0} + \kappa_2 \tau_{i,0}$, where the vectors $\kappa$ are estimated along with the other parameters. We consider similar specifications for the loadings $\alpha_j$ and the log-variances of the disturbances $\epsilon_{i,j}$. We note that the differences in education level and age in our empirical application are small.

For identification purposes we fix the first elements of $\alpha_C(E_{i,0}, \tau_{i,0})$ and $\alpha_N(E_{i,0}, \tau_{i,0})$ to be equal to one, see Cunha and Heckman (2008). These identification restrictions are sufficient for the extraction of the latent factors. We model the latent skills $F_i = (f_{i,C}, f_{i,N})'$ as random variables for which we assume the following distribution

$$F_i \sim NID(0, \Sigma_f), \quad i = 1, \ldots, N,$$  \hspace{1cm} (4.3)

where $NID(0, \Sigma_f)$ denotes the independent normal distribution with mean zero and variance matrix $\Sigma_f$. The means of the measurements in (4.1) and (4.2) are captured by $\mu_j$, for $j = C, N$. The variance matrix $\Sigma_f$ is treated as fixed and its elements may be estimated along with the other model parameters.

### 4.2.2 Childhood signals

Next to the cognitive and non-cognitive skills, the individuals also obtain additional signaling characteristics during childhood. The signals may influence potential employers, potential romantic partners and other entities, over and above the skills. We include two signals: one for education and one for a criminal record. Since the signals
are determined during childhood they are associated with the childhood cognitive and non-cognitive skills, and therefore we explicitly model them as a function of these skills (Reynolds, Temple, Robertson, & Mann, 2001).

Given that education levels summarize information for potential employers they can affect adulthood employment outcomes (Weiss, 1995). In addition, Lochner (2004) and Lochner and Moretti (2004) have argued that higher education levels increase the costs associated with crime. Similar mechanisms, over and above the effects of cognitive and non-cognitive skills, can be hypothesized for the influence of education levels on other adult outcomes such as intimate relationships, drug use and social welfare.

To incorporate the education signal, suppose that each individual chooses a level of schooling $E_i$ from the increasing set $\{1, \ldots, S\}$. We notice that the level of schooling $E_i$ can be different from the level that is pursued at the time of treatment $E_{i,0}$. However, no change in level is necessary. We model the schooling choice by a count data model, where the cognitive and non-cognitive skill factors partially determine the log intensity of the count distribution.

In particular, we specify

$$E_i \sim \text{Poisson } I_{i,s}, \quad \log I_{i,s} = X_{i,s}\beta_s + \alpha_{s,C}J_{i,C} + \alpha_{s,N}f_{i,N},$$

(4.4)

where $I_{i,s}$ is the intensity of the Poisson distribution for individual $i$, $X_{i,s}$ is a vector of observed controls and $\alpha_{s,C}$ and $\alpha_{s,N}$ are the loadings for the cognitive and non-cognitive skills$^4$.

Second, a large portion of disadvantaged youths carry over a criminal record from childhood which is likely to influence adult outcomes. Several channels have been documented for a variety of adult outcomes. For example, experimental studies have shown that the probability of employment is lower for individuals with a criminal record (Pager, 2003; Pager et al., 2009). Also, a criminal record may influence contemporaneous offending outcomes as it reduces constraints and strengthens incentives to offending (Nagin & Paternoster, 2000).

To capture the signal from a criminal record let $O_i$ denote the number of registered

$^4$A more advanced specification for the education signal would allow for a discrete choice structure with different parameters for different education levels, see for example Heckman, Stixrud, and Urzúa (2006). However, since the education levels of the sample members are generally low it is not feasible to empirically identify different parameters for different education levels.
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

offenses that are committed prior to adulthood by individual $i$. A count data model for the criminal record signal is given by

$$O_i \sim \text{Poisson} \ I_{i,o}, \quad \log I_{i,o} = X_{i,o} \beta_o + \alpha_{o,C} f_{i,C} + \alpha_{o,N} f_{i,N},$$

(4.5)

where $I_{i,o}$ is the intensity of the criminal record for individual $i$, $X_{i,o}$ captures additional explanatory variables and $\alpha_{o,C}$ and $\alpha_{o,N}$ are the loadings for the cognitive and non-cognitive skills.

It is possible to extend the model with additional signals from childhood. Given that in our framework the childhood signals are considered time-invariant the econometric difficulties for such extensions are minor. We summarize the childhood signals in the vector $G_i = (E_i, O_i)'$, for $i = 1, \ldots, N$.

4.2.3 Adult outcome model

When entering adulthood the individuals inherit two vectors that summarize the childhood skills $F_i$ and the childhood signals $G_i$. We are interested in the payoff from the childhood factors on the $M$ adult outcomes that are summarized in the vector $Y_{i,t}$. The index $t$ indicates the age of the individuals during adulthood and is arbitrarily normalized to range from $t = 1$ until $t = T$. The observation model for adulthood variable $Y_{i,t,k}$ is given by

$$Y_{i,t,k} \sim p_k(Y_{i,t,k}|\theta_{i,t,k}, \psi), \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \quad k = 1, \ldots, M,$$

(4.6)

where $p_k(Y_{i,t,k}|\theta_{i,t,k}, \psi)$ is a well-defined density function that depends on the signal $\theta_{i,t,k}$ and a vector of fixed model parameters $\psi$. We assume that $p_k(Y_{i,t,k}|\theta_{i,t,k}, \psi)$ is twice differentiable with respect to $\theta_{i,t,k}$. Note that different outcomes $k$ can be modeled by different densities $p_k()$, see also Koopman, Lucas, and Schwaab (2011), and that the signal $\theta_{i,t,k}$ is outcome specific.

---

5The age of criminal responsibility in The Netherlands is 12. Until age 18 sentencing falls under juvenile courts. In our empirical application adulthood starts at age 16, or when the individuals leave the treatment facility, and we include all registered offenses prior to this starting date.

6This is not a necessary restriction. In principle many signals can affect a single adult outcome. However by linking the signals one-to-one to the adult outcomes the interpretation for the payoffs from childhood skills and signals is simplified.
Let $\theta_{i,t} = (\theta_{i,t,1}, \ldots, \theta_{i,t,M})'$ summarize the vector of adult signals. Our analysis is based on the following dynamic linear decomposition model for the adult signals:

$$
\theta_{i,t+1} = \delta + A_t Y_{i,t} + B_t F_i + C_t G_i + D_t X_{i,t} + e_{i,t}, \quad (4.7)
$$

where $\delta$ is the mean vector, $A_t$ captures level shifts from the adult outcomes $Y_{i,t}$, $B_t$ captures the effects from the childhood cognitive and non-cognitive skills $F_i$, $C_t$ captures the effects from the childhood signals $G_i$, $D_t$ captures the effects from observed controls $X_{i,t}$ and $e_{i,t}$ is the disturbance term. The coefficient matrices $A_t$, $B_t$, $C_t$ and $D_t$ are considered fixed. Since we expect that the elements of the coefficient matrices vary smoothly with age we model them using cubic spline functions. This gives a parsimonious model specification for which we give the construction details in Appendix B.

The adult signal vector (4.7) captures persistent heterogeneity between the individuals by including the skills and signals from childhood. Further, “true” or structural effects from adult outcomes are captured by $A_t$, see also Heckman (1981a). To include unobserved persistence, or spurious correlation, we model the error term $e_{i,t}$ as a stationary vector autoregressive process of order one. In particular, we specify

$$
e_{i,t+1} = \Gamma e_{i,t} + \eta_{i,t}, \quad NID(0, \Sigma_\eta), \quad (4.8)
$$

where we restrict $\Sigma_\eta = I - \Gamma \Gamma'$ for identification purposes. The disturbance process is initialized by $e_{i,1} \sim N(0, I)$ which allows for initial deviations from the mean $\delta$. Similar models for capturing spurious correlation are considered in Keane (1994), Hyslop (1999), Heiss (2008) and Keane (2013). The main difference here is that the error term is multivariate instead of univariate. This allows the unobserved shocks $\eta_{i,t,k}$ to influence multiple future outcomes via the components of (4.7). Additionally, we follow Wooldridge (2005) and adjust the mean $\delta$ to account for persistent effects from the initial observations $Y_{i,1}$. In particular, we specify $\delta = \delta_0 + \delta_1 Y_{i,1}$, such that the effects from the initial conditions are not incorrectly attributed to $A_t$. 

84
4.2.4 Identification and estimation

Next, we discuss the estimation of the parameter vector $\psi$, which contains the unrestricted elements of the fixed matrices in equations (4.1)-(4.8). Several estimation methods are available. A two-step method first estimates the parameters that pertain to the childhood skills model (4.1)-(4.2). Given these parameter estimates some estimate for the skills $F_i$, say the posterior mean, can be obtained. In the second step the posterior mean estimate for $F_i$ is plugged in the model for the childhood signals and the adult outcomes. The parameters of these models can then be estimated using regression methods and the simulation methods of Durbin and Koopman (1997) and Shephard and Pitt (1997).

Despite the computationally attractive properties of the two-step method, Geweke and Amisano (2011) have recently argued that such methods lead to incorrect inference for the second stage model parameters and thus incorrect conclusions. Since our main interest is in these parameters we propose to estimate $\psi$ using a one-step Monte Carlo maximum likelihood method. This requires the following additional identifying assumptions for the disturbances and the distribution of the childhood skill factors.

In particular, we assume that the adult disturbance terms $\eta_{i,t}$ are independent of the disturbances $\epsilon_{j,C}$ and $\epsilon_{j,N}$, for all $i, j = 1, \ldots, N$ and $t = 1, \ldots, T$. This seems a mild assumption since the disturbances $\epsilon_{j,C}$ and $\epsilon_{j,N}$ capture measurement error from the cognitive and non-cognitive skill measures and the adult outcome disturbances are anchored to specific adult outcomes, see also Cunha et al. (2010). Further, we assume that the skill factors $F_i$ are independent of $\epsilon_{j,C}$, $\epsilon_{j,N}$ and $\eta_{l,t}$, for all $i, j, l = 1, \ldots, N$ and $t = 1, \ldots, T$. This is a standard assumption in random effects factor models, see for example Pesaran (2006, Assumption 3) and Koopman and Mesters (2014, Assumption 1b)\footnote{We emphasize that both identifying assumptions are also made in two-step procedures. However, they are generally not discussed.}

We summarize all observations in $Y = \{Y_{i,t}, Z_i, E_i, O_i\}_{i=1,\ldots,N,t=1,\ldots,T}$, where $Z_i = \{Z_{i,j}\}_{j=C,N}$. The loglikelihood for the observations is given by $\ell(\psi; Y) = \log p(Y; \psi)$, where $p(Y; \psi)$ is the joint density for the observations. For notational convenience we drop the dependence on the parameter vector $\psi$. The joint density for the observations
is given by
\[
p(Y) = \prod_{i=1}^{N} p(Y_i), \quad p(Y_i) = \int_{e_i} \int_{F_i} p(Y_i|e_i, F_i)p(e_i, F_i) \, dF_i \, de_i, \tag{4.9}
\]
where \(Y_i = \{Y_{i,t}, Z_i, E_i, O_i\}_{t=1,...,T} \) and \(e_i = (e'_{i,1}, \ldots, e'_{i,T})'\). From our assumptions above it follows that \(p(e_i, F_i) = p(e_i)p(F_i)\). To compute the likelihood contribution from individual \(i\) we need to integrate \(e_i\) and \(F_i\) from the joint density \(p(Y_i|e_i, F_i)p(e_i)p(F_i)\), where \(p(Y_i|e_i, F_i)\) is defined by the model (4.1)-(4.8). Several methods, such as particle filtering and Markov Chain Monte Carlo methods, can be used. Here we follow Mesters and Koopman (2014) and use importance sampling to facilitate the integration. Intuitively, we proceed by integrating out \(e_i\) while holding \(F_i\) fixed at its posterior modal value and we integrate out \(F_i\) while holding \(e_i\) fixed at its posterior modal value.

The joint posterior mode of \(p(e_i, F_i|Y_i)\) is defined by
\[
\{\hat{e}_i, \hat{F}_i\} = \arg \max_{e_i, F_i} p(e_i, F_i|Y_i),
\]
where we compute \(\hat{e}_i\) and \(\hat{F}_i\) using the Newton-Raphson algorithm for which we give the details in Appendix A. Using these posterior modes we rewrite the likelihood contribution for individual \(i\) as follows
\[
p(Y_i) = \int_{e_i} \int_{F_i} \frac{p(Y_i|e_i, F_i)p(e_i)p(F_i)}{g(F_i|Y_i; \hat{e}_i)g(e_i|Y_i; \hat{F}_i)} \, g(F_i|Y_i; \hat{e}_i) \, g(e_i|Y_i; \hat{F}_i) \, dF_i \, de_i, \tag{4.10}
\]
where \(g(F_i|Y_i; \hat{e}_i)\) and \(g(e_i|Y_i; \hat{F}_i)\) are the importance densities. In practice we choose \(g(F_i|Y_i; \hat{e}_i)\) and \(g(e_i|Y_i; \hat{F}_i)\) to follow Gaussian distributions for which we discuss the details in Appendix A. We notice that \(g(F_i|Y_i; \hat{e}_i)\) considers \(e_i\) fixed and \(g(e_i|Y_i; \hat{F}_i)\) considers \(F_i\) fixed. This separates the integration of \(F_i\) and \(e_i\).

Further, when we restrict \(p(e_i) \equiv g(e_i)\) and \(p(F_i) \equiv g(F_i)\), such that the marginal distributions of the disturbances and the error terms of the original model \(p\) and the
importance sampling model \( g \) are identical, it follows from Bayes’ rule that

\[
p(\mathcal{Y}_i) = g(\mathcal{Y}_i; \hat{e}_i)g(F_i; \hat{F}_i) \int_{e_i} \int_{F_i} \frac{p(\mathcal{Y}_i|e_i, F_i)}{g(\mathcal{Y}_i|F_i; \hat{e}_i)g(F_i|\mathcal{Y}_i; \hat{e}_i)g(e_i|\mathcal{Y}_i; F_i)} dF_i de_i, \tag{4.11}
\]

where \( g(\mathcal{Y}_i; \hat{e}_i) \) and \( g(\mathcal{Y}_i; \hat{F}_i) \) can be viewed as the likelihoods of the models implied by the importance sampling densities, which are re-weighted by the weights

\[
w_i = \int_{e_i} \int_{F_i} \frac{p(\mathcal{Y}_i|e_i, F_i)}{g(\mathcal{Y}_i|F_i; \hat{e}_i)g(F_i|\mathcal{Y}_i; \hat{e}_i)g(e_i|\mathcal{Y}_i; F_i)} dF_i de_i,
\]

to correct for the fact that \( g(\mathcal{Y}_i; \hat{e}_i)g(\mathcal{Y}_i; \hat{F}_i) \neq p(\mathcal{Y}_i) \).

We approximate the weights by simulation methods since no analytical solution is available when \( p(\mathcal{Y}_i|e_i, F_i) \) is non-Gaussian, see Durbin and Koopman (2012, Part 2) for a more elaborate discussion. The Monte Carlo estimate for the weights is given by

\[
\hat{w}_i = L^{-1} \sum_{l=1}^{L} \frac{p(\mathcal{Y}_i|e_i^{(l)}, F_i^{(l)})}{g(\mathcal{Y}_i|F_i^{(l)}; \hat{e}_i)g(F_i^{(l)}|\mathcal{Y}_i; \hat{e}_i)g(e_i^{(l)}|\mathcal{Y}_i; F_i^{(l)})}, \tag{4.12}
\]

where the samples \( e_i^{(l)} \) and \( F_i^{(l)} \) are drawn from \( g(e_i|\mathcal{Y}_i; \hat{F}_i) \) and \( g(F_i|\mathcal{Y}_i; \hat{e}_i) \), for \( L = 1, \ldots, L \). When replacing the weights by their Monte Carlo approximation it follows that the likelihood contribution for individual \( i \) is given by

\[
\hat{p}(\mathcal{Y}_i) = g(\mathcal{Y}_i; \hat{e}_i)g(\mathcal{Y}_i; \hat{F}_i)\hat{w}_i, \tag{4.13}
\]

where it follows from Geweke (1989) that \( \hat{p}(\mathcal{Y}_i) \to p(\mathcal{Y}_i) \) as \( L \to \infty \) if the variance of \( \hat{w}_i \) is finite.

The key insight for the importance sampling approach proposed here is that the likelihood contribution of individual \( i \) can be approximated using Monte Carlo methods independent from the likelihood contribution of individual \( j \), for all \( i, j = 1, \ldots, N \). This implies that per individual the number of random components that need to be integrate from the likelihood is small, i.e. only \( F_i \) and \( e_i \). The result is that we can construct highly accurate importance densities such that the weights \( \hat{w}_i \) are very close to one. The technical details for implementation are given in Appendix A and on our website we provide the implementation code.
4.3 Sample description and data origins

We use data from the NSCR 17up study, a longitudinal study that follows disadvantaged youths into adulthood. The individuals that we consider had all been institutionalized in a juvenile treatment center in the early 1990s. We only include individuals who stayed in the institution for more than two months, had a complete treatment file and were not being treated for sexual offending behavior. The original sample of 270 males includes all males that were discharged from the treatment facility between January 1989 and June 1996. The original 270 females were discharged between January 1990 and March 1999.

In The Netherlands juveniles are sent to treatment institutions for a variety of reasons. Examples include serious behavioral problems, criminal activity and disrupted family situations. Often a combination of these three aspects makes it impossible for the individuals to remain at home. The ages in the institution range between 10 and 20, where the median is just above 16. According to the prevailing law, from the age of 12 treatment can be imposed as a criminal law measure. Before the age of 12 treatment can only be imposed as a civil law measure. Eighty percent of our sample was sent to the treatment institution under a civil law measure. The distinction between criminal and civil law says little about the severity of behavioral problems or whether an individual has a conviction prior to treatment in the institution (Wijkman et al., 2006). While in the institution, behavioral problems are treated and low-level education is provided. The original sample is not representative for the general population of institutionalized males and females as their development problems were found more severe when compared to those from other juvenile treatment centers (Boendermaker, 1999). The sample that we study may be considered as a highly disadvantaged group of juveniles.

Face-to-face interviews were conducted with a subsample of the original 540 males and females. At the start of the interview phase in 2010, 22 of the original individuals had died, 14 had emigrated, 5 were living in institutions and another 19 could not be traced. The remaining 499 males and females were approached for interviews. Out of the 499 individuals 116 males and 132 females completed a full interview, after giving informed consent. A response analysis was conducted to verify the representativeness of the interviewed versus the original sample. Besides individuals without a regular
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

place of living the subsample can be regarded as representative for the original sample. Further details for the interviews and the response analysis are given in van der Geest et al. (2013). In the next sections we detail the origins of the data used and the construction of the variables.

4.3.1 Treatment files

Measurements that are related to the cognitive and non-cognitive skills are obtained from the individuals’ treatment files, which were obtained from the archive of the juvenile treatment facility. These files generally contained the results from psychological and psychiatric tests, advisory notes on extensions, and treatment evaluations. The reports in the files are in most cases prepared by forensic psychologists and psychiatrists. However, also external reports from the Dutch Child Protection Agency and other organizations responsible for the supervision of juveniles were found. For all individuals progress reports regarding their treatment were compiled by a multi-disciplinary team.

Although the contents of the files varied between individuals, we were able to extract a large number of common items. A complete list of measurements is given in Table 4.1. These measurements are included in the cognitive and non-cognitive skill vectors $Z_{i,C}$ and $Z_{i,N}$. For the cognitive skills the intelligence scores were measured using the Wechsler Intelligence Scale for Children (revised for The Netherlands) and the Raven Progressive Matrices. These were then turned into categories according to prevailing norm values that ranged between retarded and highly gifted. In our sample almost 10% of the 116 males and 8% of the 132 females were considered retarded, whereas none were considered highly gifted. The females score slightly higher on the intelligence scale, when compared to the males. In addition to the intelligence score we also include measurements for cognitive improvement during treatment and perspective for future education to extract the cognitive skill factor. The measurements for cognitive improvements and future education were assessed by psychologists.

For the non-cognitive skills we have obtained measurements for a large number of personality traits. The variables for neuroticism, impulsiveness, thrill seeking and extroversion were derived from standard questionnaires used in the files and validated self reports. Relevant guidelines, such as the ATL (Adolescent Temperament List) and the NPV-J (Netherlands Personality Questionnaire - Youth) were used to construct
the norm values. Unfortunately only the norm values are available to us and we cannot use the original questions to reconstruct the scores. Additionally, we included a large number of variables that were constructed by the treating psychologists. Examples include measurements for aggressive behavior, anti-social behavior, conscience development, authority problems, social skills, suicide attempts, depression and some indications whether the treatment was successful. We emphasize that all measurements are coded monotonically increasing, such that higher scores on the skills imply more (non-)cognitive skills. We refer to van der Geest and Bijleveld (2008) for additional discussion of these measurements.

Additional variables that we obtained from the treatment files include general personal characteristics, such as ethnicity, dates of admission and discharge, and the education level that the individuals were pursuing while staying in the treatment facility. We coded the education levels according to the ISCED classification (International Standard Classification of Education), which was adjusted to differentiate between vocational and non-vocational education in The Netherlands in the European Value Studies (2013). The classification is summarized in Table 4.2. The childhood skills factor model in Section 4.2.1 corrects the model coefficients for this educational level and age at the time of intake. In the childhood signals model (4.4)-(4.5) we also include control variables for ethnicity and the length of stay in the institution.

4.3.2 Interview data and register data

Between 2010 and 2012 a subsample of the original sample were interviewed. The interview consisted of two parts from which we obtained data. First, a structured questionnaire was filled in which covered a wide array of topics. For the purpose of this study we only use the questions that were related to education, from which we construct the education signal $E_i$. In particular, the individuals were asked whether they completed a degree while staying in the juvenile treatment facility and whether they had obtained any additional degrees after leaving the facility. We coded the highest level of education for which a degree was obtained using the ISCED classification given in Table 4.2.

In panels i.a and i.b of Figure 4.1 we show the frequencies for the education levels for the males and females. We find that the majority of the females did not complete any
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

<table>
<thead>
<tr>
<th>Measurement</th>
<th>MALES (# = 116)</th>
<th>FEMALES (# = 132)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligence</td>
<td>0 2.710 0.862</td>
<td>2.820 0.939</td>
</tr>
<tr>
<td>Perspective for future schooling</td>
<td>1 3.304 0.790</td>
<td>3.021 0.829</td>
</tr>
<tr>
<td>Cognitive improvements after treatment</td>
<td>2 3.044 0.710</td>
<td>2.875 1.053</td>
</tr>
<tr>
<td><strong>Non-Cognitive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurotic behavior</td>
<td>0 1.740 0.783</td>
<td>1.696 0.847</td>
</tr>
<tr>
<td>Impulsive behavior</td>
<td>1 1.832 0.878</td>
<td>1.844 0.729</td>
</tr>
<tr>
<td>Thrill seeking</td>
<td>2 2.365 0.795</td>
<td>1.640 0.714</td>
</tr>
<tr>
<td>Conscience development</td>
<td>3 2.255 0.675</td>
<td>1.753 0.682</td>
</tr>
<tr>
<td>Self image</td>
<td>4 1.939 0.708</td>
<td>1.748 0.611</td>
</tr>
<tr>
<td>Aggressive behavior</td>
<td>5 2.105 0.985</td>
<td>2.118 0.944</td>
</tr>
<tr>
<td>Anti-social behavior</td>
<td>6 2.765 0.818</td>
<td>2.202 0.829</td>
</tr>
<tr>
<td>Authority problems</td>
<td>7 2.440 1.134</td>
<td>2.315 0.874</td>
</tr>
<tr>
<td>Social skills</td>
<td>8 2.139 0.856</td>
<td>1.746 0.758</td>
</tr>
<tr>
<td>Depression</td>
<td>9 3.248 1.085</td>
<td>2.600 0.863</td>
</tr>
<tr>
<td>Self-destruction</td>
<td>10 2.861 0.423</td>
<td>2.305 0.863</td>
</tr>
<tr>
<td>Suicide attempt</td>
<td>11 2.911 0.318</td>
<td>2.674 0.640</td>
</tr>
<tr>
<td>Relationship with peers</td>
<td>12 2.248 0.883</td>
<td>1.793 0.691</td>
</tr>
<tr>
<td>Individual psychotherapy</td>
<td>13 1.867 0.340</td>
<td>1.621 0.485</td>
</tr>
<tr>
<td>Motivation to be treated</td>
<td>14 3.088 1.013</td>
<td>3.226 0.905</td>
</tr>
<tr>
<td>Acknowledgment crimes</td>
<td>15 3.149 0.558</td>
<td>3.828 0.833</td>
</tr>
<tr>
<td>Understanding of risks</td>
<td>16 3.013 0.733</td>
<td>3.243 0.917</td>
</tr>
<tr>
<td>Understanding of consequences of actions</td>
<td>17 3.033 0.814</td>
<td>3.275 0.987</td>
</tr>
<tr>
<td>Impulsive behavior AT</td>
<td>18 3.120 0.805</td>
<td>3.167 0.799</td>
</tr>
<tr>
<td>Change impulsive behavior AT</td>
<td>19 1.707 0.543</td>
<td>1.297 0.457</td>
</tr>
<tr>
<td>Change aggressive behavior AT</td>
<td>20 1.663 0.537</td>
<td>1.382 0.551</td>
</tr>
<tr>
<td>Change social skills T</td>
<td>21 1.692 0.485</td>
<td>1.299 0.458</td>
</tr>
<tr>
<td>Emphatic behavior AT</td>
<td>22 2.769 0.681</td>
<td>3.019 0.739</td>
</tr>
<tr>
<td>Relationship staff AT</td>
<td>23 3.207 0.801</td>
<td>2.984 0.830</td>
</tr>
<tr>
<td>Relationship Peers AT</td>
<td>24 3.228 0.644</td>
<td>2.849 0.656</td>
</tr>
<tr>
<td>Success treatment</td>
<td>25 3.196 0.924</td>
<td>3.539 0.759</td>
</tr>
<tr>
<td>Probability recidivism</td>
<td>26 2.978 1.033</td>
<td>2.814 1.062</td>
</tr>
</tbody>
</table>

Table 4.1: Summary statistics for measurements for the extraction of cognitive and non-cognitive skill factors. AT refers to after treatment and T refers to treatment. All variables are measured on a 4 or 5 points scale.
<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>primary education not completed</td>
</tr>
<tr>
<td>1</td>
<td>primary education (special)</td>
</tr>
<tr>
<td>2</td>
<td>primary education</td>
</tr>
<tr>
<td>3</td>
<td>secondary vocational low (LTS/VBO)</td>
</tr>
<tr>
<td>4</td>
<td>basic secondary school (MAVO/MULU)</td>
</tr>
<tr>
<td>5</td>
<td>secondary vocational (MBO/MTS/leerwezen &lt; 2 years)</td>
</tr>
<tr>
<td>6</td>
<td>secondary vocational (MBO/MTS/leerwezen 2&lt;5 years)</td>
</tr>
<tr>
<td>7</td>
<td>medium secondary school (HAVO)</td>
</tr>
<tr>
<td>8</td>
<td>higher secondary school (VWO)</td>
</tr>
<tr>
<td>10</td>
<td>tertiary education (HBO)</td>
</tr>
<tr>
<td>11</td>
<td>tertiary education (WO, bachelor)</td>
</tr>
</tbody>
</table>

Table 4.2: Education levels for classifying education in The Netherlands. We added the category for special primary education to distinguish between regular and special primary school, where special education is provided for those with mental or physical disabilities. We further omitted higher categories for master level university education and PhD programs as these are not attained within our sample.

education above primary school level. For the males the majority obtained low-level vocational education. We point out that for males the treatment facility nearly always provided a low level vocational degree, whereas for females very few left the treatment facility with any degree. Further, for both males and females secondary vocational level training of less than 2 years is the most commonly attained level. This category includes degrees for hairdressing and nail styling, which are frequently obtained by the females in our sample. Also included are cooking school degrees and construction degrees which are popular among the males. Degrees higher than level 4 are typically obtained after leaving the juvenile treatment facility. The vast majority of these are part-time degrees, which typically include only one weekday of actual schooling and where the remaining four days are spent working in a relevant job.

The second part of the interview included filling in a life-history calendar. This tool aims to retrospectively reconstruct the life of the individual for a variety of adult life domains (Caspi et al., 1996). In Figure 4.2 we provide an example of the calendar used. For the purpose of this study we used the information related to drug use and intimate relationships. The adulthood drug use variable was coded as a binary variable which was set equal to one whenever the individual indicated the use of hard
drugs (cocaine, heroine, speed, amphetamine and methadon) in a specific age-period. Similarly, intimate relationships were coded as binary variables being equal to one whenever the individual indicated to be in an intimate relationship in a specific age-period. In the next section we present the descriptive statistics for all adult outcome variables.

In addition to the self-reported data from the interviews we collected official register information. Two sources of register data were used: judicial documentation and centralized employment records. The judicial documentation contains abstracts of The Netherlands Ministry of Justice. These are comparable to rap sheets in the US. The abstracts contain information on every case that is sent to the Public Prosecutor’s Office and the verdict that follows from it. We select only those cases that did not end in an acquittal or dismissal for a lack of evidence. The abstracts also contain information on the date and the type of the offense. The abstracts are available for each individual from age 12 and onwards, 12 being the age of criminal responsibility.

We use the rap sheets for three purposes. First, we counted for each individual all offenses that were committed before age 16\(^8\). This count was used to construct the criminal records signal \(O_i\). In panels ii.a and ii.b of Figure 4.1 we show the distributions of the number of prior registered offenses for both males and females. The number of prior offenses is often larger for males when compared to females. The maximum number of prior offenses that we find is 25 for males and 24 for females. Second, we use the rap sheets to code the adulthood crime variable. For this we consider all serious offenses following the definition given in Loeber et al. (1998)\(^9\). A binary outcome variable was constructed to be equal to one whenever at least one serious offense was committed in a specific age period. Third, we use the rap sheets to construct an incarceration variable. Since sentences in The Netherlands are often short we code the incarceration variable as the percentage of days in a year that an individual was not incarcerated. This “exposure” variable is used as a control variable in the adult outcome model (4.7), i.e. included in \(X_{i,t}\).

\(^8\)In our empirical application 16 is the age of entering adulthood. If the individuals were still in the treatment facility after this age we also included the offenses that were committed during the remainder of their stay in the facility. There were few registered offenses obtained while institutionalized.

\(^9\)This classification includes all violent offenses, felony larceny, auto theft, burglary, breaking and entering, carjacking, forgery and counterfeiting, fraud, dealing in stolen property, embezzlement, drug trafficking, arson, weapons and firearms violations.
Employment and social welfare data is obtained from the Ministry of Social Affairs. The information consists of individual-level employment and social welfare histories. For each employment spell we know the exact start and ending date of the contract. Whether a position was full-time or part-time remains unknown to us as we have no information on the exact number of hours spent working. We construct the employment variable to include all spells that pertain to regular employment, hereby excluding employment spells that pertain to employment via a temporary job agency. The latter distinction is shown relevant in van der Geest et al. (2011) and Mesters et al. (2014) since employment through a temporary job agency in the Netherlands often is seasonal or project based, and generally lasts for short spells of a few weeks providing little long-term prospects compared with regular jobs. We code the employment variable as a binary indicator, which we set equal to one whenever more than 3 months of regular employment were found within one age year. Some additional experiments have shown that the three months indicator is not sensitive for the results, see also van den Berg, Mesters, Bijleveld, and Hendriks (2014) where this is investigated in more detail.

Next to the employment outcomes, three types of social welfare are recovered from the SZW database: unemployment insurance, disability insurance and public assistance. The majority of the social welfare benefits are public assistance benefits. These are unconditional cash transfers for which only financial need needs to be proven. Typically these spells last for long periods often consisting of multiple years. A small part of the welfare benefits consists of unemployment and disability insurance. For these the requirements include a substantial period (at least 6 months) of previous employment, and for disability insurance a proof of illness from a doctor. Since all forms of welfare provide financial gains for the recipients we take all together and define the total social welfare outcome as equal to one whenever the any form of social welfare was claimed during a particular year.

### 4.3.3 Summary statistics adulthood

Next, we discuss the descriptive statistics for the adult outcomes variables ($N = 116$ males and $N = 132$ females), which include binary indicators for serious offending, employment, social welfare, drug use and intimate relationships. Figure 4.3 shows the average rates of all outcome variables for different age-periods.
For 116 males we have 1746 observational periods between ages 16 and 32 and in 24.1% of these periods a serious offense was committed. The majority of the offenses are committed between ages 16 and 20, after which the average rate declines consistently. This “age-crime” curve has its familiar shape, but the average rate is much higher when compared to the general population (Farrington, 1986; van der Geest et al., 2009). For females a similar declining pattern in the serious offending rate is observed, but the average rate of offending is much lower at 8.2% out of 1912 observational periods for the 132 females.\footnote{The differences in the number of observed periods between males and females are partially due to differences in the ages that they left the treatment facility. Males left the institution on average around age 17.376 (sd 1.384) and females at age 16.576 (sd 1.279). And there are more females when compared to males in the samples.}

The employment rate increases steadily for both males and females from around
zero at age 16 to about 35% around age 25. After age 25 the rates for both the males and females remain stable. Further inspection shows that there is a small subsample of individuals who have steady employment from around age 22 onwards. The majority of individuals have infrequent employment spells and for another small subsample no registered employment spells were found. Similar employment participation rates are found for other samples of disadvantaged youths in Grogger (1995) and Levitt and Venkatesh (2001).

The social welfare rate starts to increase from age 22 onwards. For females the rate reaches its maximum by age 28 at around 40%. The difference between males and females is quite substantial as the female rate exceeds the male rate for the entire observational period. We note that welfare participation rates are much higher for our sample as compared to the general population (e.g., Mesters et al., 2014).

Drug use in our sample is high. For males the rate is 0.266 in the observed per individual per age-periods. Recall that we only included hard drugs, such as cocaine and heroine, in the outcome variable. The distribution over the ages is steady, showing
only a modest decline with maturation. Around age 16 the rate is approximately the same for males and females. For females however, the rate declines steadily from age 16 onwards. At age 32 the rate is below 0.1.

The self-reported relationship indicator shows that on average approximately 50% of the males and 70% of the females were in an intimate relationship at some point during the observational period. Often several partners were reported between ages 16 and 32. Approximately 30% of the females was in a relationship at age 16, whereas only 10% of the males was in a relationship at this age. The female rate increases rapidly to 70% at age 18 after which it remains stable. For males a more gradual increase in the rate is found, which almost matches the female rate at the end of the observation period.

The overall correlations between the outcome variables are given in Table 4.3. We find that for both males and females serious offending is negatively correlated with employment, social welfare and intimate relationships and positively correlated with drug use. Both employment and social welfare are positively associated with intimate relationships and negatively with drug use. Finally, drug use and intimate relationships are negatively related. Overall the influence of drug use seems more important for females, whereas employment seems more influential for males.

From a statistical point of view these correlations can be structural, in the sense that for example employment leads to lower offending, or spurious in the sense that underlying skills or signals (or other variables) generate the correlations. In the next section we formally decompose these aspects using our statistical model given in Section 4.2.

### 4.4 Empirical estimation results

In this section we discuss our empirical results for the samples of disadvantaged males and females discussed in Section 4.3. We consider the statistical model given in Section 4.2. We estimate the model parameters separately for the samples of $N = 116$ males and $N = 132$ females. We use $L = 100$ draws from the importance sampling densities to evaluate the Monte Carlo likelihood in (4.13), which is optimized with respect to the parameters using numerical methods (Nocedal & Wright, 1999). Section 4.4.1 discusses the parameter estimates for the case where the payoffs from the childhood
Table 4.3: Overall correlations between the adulthood outcome variables for males and females.

<table>
<thead>
<tr>
<th></th>
<th>SO</th>
<th>EM</th>
<th>SW</th>
<th>DR</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serious offending</td>
<td>SO</td>
<td>-0.207</td>
<td>-0.024</td>
<td>0.159</td>
<td>-0.050</td>
</tr>
<tr>
<td>Regular employment</td>
<td>EM</td>
<td>-0.207</td>
<td>-0.001</td>
<td>-0.022</td>
<td>0.146</td>
</tr>
<tr>
<td>Social welfare</td>
<td>SW</td>
<td>-0.024</td>
<td>-0.001</td>
<td>-0.004</td>
<td>0.087</td>
</tr>
<tr>
<td>Drug use</td>
<td>DR</td>
<td>0.159</td>
<td>-0.022</td>
<td>-0.004</td>
<td>-0.055</td>
</tr>
<tr>
<td>Intimate relationships</td>
<td>IR</td>
<td>-0.050</td>
<td>0.146</td>
<td>0.087</td>
<td>-0.055</td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serious offending</td>
<td>SO</td>
<td>-0.116</td>
<td>-0.042</td>
<td>0.181</td>
<td>-0.015</td>
</tr>
<tr>
<td>Regular employment</td>
<td>EM</td>
<td>-0.116</td>
<td>-0.028</td>
<td>-0.096</td>
<td>0.112</td>
</tr>
<tr>
<td>Social welfare</td>
<td>SW</td>
<td>-0.042</td>
<td>0.028</td>
<td>-0.073</td>
<td>0.038</td>
</tr>
<tr>
<td>Drug use</td>
<td>DR</td>
<td>0.181</td>
<td>-0.096</td>
<td>-0.073</td>
<td>-0.110</td>
</tr>
<tr>
<td>Intimate relationships</td>
<td>IR</td>
<td>-0.015</td>
<td>0.112</td>
<td>0.038</td>
<td>-0.110</td>
</tr>
</tbody>
</table>

skills, signals and contagion in adult outcomes are considered age-invariant, i.e. $A_t = A$, $B_t = B$, $C_t = C$ and $D_t = D$, for all $t = 1, \ldots, T$. This gives the overall importance and significance for the different model components. In Section 4.4.2 we study the age-varying effects of the childhood skills and signals on the adult outcomes and the age-variation in the adult outcome contagion effects.

### 4.4.1 Age-invariant effects

The main parameter estimation results are presented in Tables 4.4 to 4.6. In Table 4.4 we show the parameter estimates for dynamic contagion in the adult outcomes (matrix $A$) and in Table 4.5 we show the parameter estimates for the effects of the cognitive and non-cognitive skills. We include the estimates for the adult outcomes (matrix $B$) and the childhood signals ($\alpha_{s,C}$, $\alpha_{s,N}$, $\alpha_{o,C}$ and $\alpha_{o,N}$). In Table 4.6 we show the effects of the childhood signals on the adult outcomes (matrix $C$).

We discuss the results for males and females separately. For males we find that cognitive skills reduce the probability of serious offending during adulthood (-0.209). Also, they increase the adulthood employment probability (0.267). For the other adulthood outcome variables and childhood signals we find no significant effects for the cognitive
skills. Non-cognitive skills are overall more important for the males when compared to cognitive skills. In particular, non-cognitive skills significantly reduce the intensity of the criminal record before age 16. Additionally, non-cognitive skills have lasting effects on adulthood serious offending (-0.336) and regular employment (0.349). Thus, both cognitive and non-cognitive skills explain the adult serious offending and employment outcomes, but the magnitude of the coefficients is larger for the non-cognitive skills.

Additional effects from the childhood signals indicate that education levels reduce serious offending during adulthood over and above the effects of the childhood skills (-0.113). This points to increasing costs associated with crime for higher education levels (Lochner, 2004). The criminal record signal increases the serious offending probability (0.034), the social welfare probability (0.044) and, perhaps surprisingly, the probability for intimate relationships (0.030). The substantial lasting effects from the criminal record suggest that this “mark”, that is obtained before age 16, severely impacts adult
outcomes, see also the discussion in Lemert (1967).

The effects from dynamic contagion in adulthood outcomes (Table 4.4) indicate that all outcomes are dependent on their own previous outcomes in a structural way. This form of state dependence implies that for example serious offending significantly increases the probability for serious offending in the next period. We find large state dependence coefficients for employment, drug use and intimate relationships. The state dependence effects can be regarded as positive for employment and intimate relationships, or negative as for drug use and social welfare participation.

In addition, there are multiple significant cross-dependencies. Serious offending is significantly reduced by employment and social welfare, whereas drug use significantly increases the serious offending probability. Employment probabilities are significantly decreased by previous offending and drug use and increased by intimate relationships. Finally, the probability for social welfare participation is significantly decreased by serious offending and employment. Drug use and intimate relationships are less influenced by the other adulthood outcomes.

Summarizing, for males we find a mixture of influences from childhood skills and signals, and dynamic contagion in adult outcomes. Employment and social welfare are negatively linked with serious offending during adulthood, and employment and serious offending are predictable by the cognitive and non-cognitive skills. The adulthood outcomes for drug use and intimate relationships are harder to predict using childhood factors but influence employment and serious offending during adulthood.

For females cognitive skills have no significant effects on the childhood signals, nor on the adult outcomes. The non-cognitive skills significantly explain the education level and the criminal record that is obtained before age 16. In particular, the intensity of the education signal is significantly increased by the non-cognitive skills (0.105). This effect is much stronger when compared to the males. The intensity of the criminal record is reduced when non-cognitive skills are higher (-0.373). For the adulthood outcome variables we find similar results from non-cognitive skills as for males. Non-cognitive skills reduce serious offending (-0.314) and increase employment (0.235).

Childhood signals have more lasting impacts on the adulthood outcomes for females. In particular, education level increases the probability of regular employment (0.168) and decreases the probability of drug use (-0.308). This indicates that females can gain much from increasing their education level. Even more so than males. The
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

Males
(N = 116)

<table>
<thead>
<tr>
<th>Variable</th>
<th>(SO_{t-1})</th>
<th>(EM_{t-1})</th>
<th>(SW_{t-1})</th>
<th>(DR_{t-1})</th>
<th>(IR_{t-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SO_t)</td>
<td>0.759** 0.172</td>
<td>-1.393** 0.353</td>
<td>-0.394* 0.240</td>
<td>0.913** 0.218</td>
<td>-0.297 0.169</td>
</tr>
<tr>
<td>(EM_t)</td>
<td>-0.588** 0.283</td>
<td>3.284** 0.240</td>
<td>-0.016 0.274</td>
<td>-0.565** 0.272</td>
<td>0.597** 0.194</td>
</tr>
<tr>
<td>(SW_t)</td>
<td>-0.474* 0.274</td>
<td>-0.687** 0.269</td>
<td>2.774** 0.257</td>
<td>-0.148 0.286</td>
<td>0.299 0.220</td>
</tr>
<tr>
<td>(DR_t)</td>
<td>0.221 0.313</td>
<td>0.269 0.309</td>
<td>-0.291 0.376</td>
<td>5.818** 0.324</td>
<td>-0.145 0.261</td>
</tr>
<tr>
<td>(IR_t)</td>
<td>-0.354 0.216</td>
<td>0.198 0.203</td>
<td>0.293 0.250</td>
<td>-0.294 0.228</td>
<td>3.972** 0.268</td>
</tr>
</tbody>
</table>

Females
(N = 132)

<table>
<thead>
<tr>
<th>Variable</th>
<th>(SO_{t-1})</th>
<th>(EM_{t-1})</th>
<th>(SW_{t-1})</th>
<th>(DR_{t-1})</th>
<th>(IR_{t-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SO_t)</td>
<td>1.141** 0.257</td>
<td>-1.302** 0.412</td>
<td>-0.433** 0.260</td>
<td>0.713** 0.274</td>
<td>0.150 0.226</td>
</tr>
<tr>
<td>(EM_t)</td>
<td>-2.088** 0.658</td>
<td>2.716** 0.191</td>
<td>0.221 0.197</td>
<td>-0.194 0.284</td>
<td>0.299 0.194</td>
</tr>
<tr>
<td>(SW_t)</td>
<td>-0.665* 0.380</td>
<td>-0.421* 0.226</td>
<td>3.819** 0.218</td>
<td>-0.918** 0.289</td>
<td>-0.831** 0.196</td>
</tr>
<tr>
<td>(DR_t)</td>
<td>-0.026 0.392</td>
<td>-0.443 0.315</td>
<td>-0.689** 0.285</td>
<td>4.743** 0.369</td>
<td>-0.514** 0.238</td>
</tr>
<tr>
<td>(IR_t)</td>
<td>-0.343 0.285</td>
<td>0.164 0.200</td>
<td>0.089 0.186</td>
<td>-0.972** 0.224</td>
<td>2.684** 0.163</td>
</tr>
</tbody>
</table>

Table 4.4: Structural effects from dynamic contagion in adulthood outcomes (matrix A). The labels refer to: \(SO\) serious offending, \(EM\) employment, \(SW\) social welfare, \(DR\) drugs use and \(IR\) intimate relationships. The indication * implies significance at the \(\alpha = 0.1\) level and ** implies significance at the \(\alpha = 0.05\) level.

Criminal records signal increases the probability of adulthood serious offending (0.098) and reduces the probability of regular employment (-0.080). Thus, as well as for males, having a criminal record before age 16 has severe consequences.

When compared to the males we find similar patterns for the adulthood contagion effects for females. The probability of serious offending is reduced by employment and social welfare, whereas drug use increases the probability of serious offending. The employment probability decreases with serious offending and social welfare participation decreases with drug use and intimate relationships. Interestingly, for females there is strong negative interaction between drug use and intimate relationships. Drug use is further reduced by social welfare participation.

In sum, for females cognitive skills have less impact on the observed adulthood outcomes. Instead, non-cognitive skills and childhood signals, which are driven by the childhood non-cognitive skills, and dynamic influences between adult outcomes have more explanatory power.
### Table 4.5: Age-invariant effects from cognitive and non-cognitive skills on childhood signals and adulthood socioeconomic outcomes (loadings $\alpha_{s,C}$, $\alpha_{s,N}$, $\alpha_{o,C}$ and $\alpha_{o,N}$ and matrix $B$). The indication $^*$ implies significance at the $\alpha = 0.1$ level and $^{**}$ implies significance at the $\alpha = 0.05$ level.

<table>
<thead>
<tr>
<th>Childhood signal</th>
<th>Males ($N = 116$)</th>
<th>Females ($N = 132$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cognitive</td>
<td>Non-Cognitive</td>
</tr>
<tr>
<td>Education level</td>
<td>0.015 0.067</td>
<td>0.021 0.079</td>
</tr>
<tr>
<td>Criminal record</td>
<td>0.105 0.072</td>
<td>-0.330** 0.068</td>
</tr>
</tbody>
</table>

### Table 4.6: Age-invariant effects from childhood signals on adulthood socioeconomic outcomes (matrix $C$). The indication $^*$ implies significance at the $\alpha = 0.1$ level and $^{**}$ implies significance at the $\alpha = 0.05$ level.

<table>
<thead>
<tr>
<th>Adulthood</th>
<th>Males ($N = 116$)</th>
<th>Females ($N = 132$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Education level</td>
<td>Criminal record</td>
</tr>
<tr>
<td>Serious offending</td>
<td>-0.113* 0.068</td>
<td>0.034** 0.015</td>
</tr>
<tr>
<td>Regular employment</td>
<td>0.086 0.066</td>
<td>0.012 0.023</td>
</tr>
<tr>
<td>Social welfare</td>
<td>-0.129 0.090</td>
<td>0.044** 0.022</td>
</tr>
<tr>
<td>Drug use</td>
<td>-0.034 0.085</td>
<td>0.015 0.027</td>
</tr>
<tr>
<td>Intimate relationships</td>
<td>-0.015 0.055</td>
<td>0.030* 0.018</td>
</tr>
</tbody>
</table>

102
4.4.2 Age-varying effects

In this section we discuss the age-varying estimates $A_t$, $B_t$ and $C_t$. These estimates allow us to study how the payoffs from childhood skills, signals and adulthood vary with age during adulthood. For males the estimates are summarized in Figures 4.4, 4.6 and 4.8, whereas for females the estimates are summarized in Figures 4.5, 4.7 and 4.9.

We first discuss the estimates for the males. Figure 4.6 shows two significant age-varying effects for the cognitive skills. First, cognitive skills become important predictors for employment from age 23 onwards. This is understandable as employment before age 23 is likely to be rather unskilled, whereas with maturation cognitive skills become more important for employment. Second, cognitive skills become important for increasing the probability of intimate relationships after age 28. The non-cognitive skills are equally important in increasing the employment probability from age 22-23 onwards. Thus, for employment both cognitive and non-cognitive skills are important at later stages in life. Non-cognitive skills further significantly reduce the offending probability between ages 16 and 18. While the estimate remains negative after age 18 it is no longer significant.

In Figure 4.8 we show the age-varying payoffs from childhood signals for males. The education level signal significantly reduces serious offending from age 23 onwards. For employment we find mixed effects from the education signal. Until age 22 the education level significantly reduces the employment probability, while after age 24 it significantly increases the employment probability. Further, the education signal significantly reduces the probability for social welfare. Thus, education provides a barrier over and above the childhood skills to reduce social welfare participation. The criminal records signal has lasting effects on serious offending. This can be due to identity transformations in the sense that the criminal records makes individuals see themselves as a criminal (e.g., Lemert, 1967). The criminal records signal further significantly increases the probability of social welfare from age 23 onwards.

In Figure 4.4 we show the age-varying effects from dynamic contagion in the adult outcomes. We only discuss a few findings. First, employment significantly reduces serious offending from age 22 onwards. This is consistent with a large literature in criminology that documents that employment during adolescence increases offending
and only employment during adulthood reduces offending (e.g., Shover, 1996; Uggen, 2000; Paternoster et al., 2003). Second, intimate relationships significantly increase the employment probability after age 26. Third and finally, employment reduces social welfare participation at later ages.

Next, we discuss the estimates for the females. In Figure 4.7 we show the age-varying effects for cognitive and non-cognitive skills on adulthood outcomes. For the entire observational period we find no significant effects for the cognitive skills. Non-cognitive skills significantly increase the employment probability until age 24 after which the estimate converges to zero. This pattern is reversed when compared to the males, for whom the non-cognitive skills increased employment after age 23. Non-cognitive skills further significantly lower the probability for intimate relationships between ages 16 and 19.

Figure 4.9 shows the age-varying estimates for childhood signals. The education level significantly reduces serious offending after age 23. Around the same age it significantly increases the employment probability. Additionally, the education signal reduces the probability of social welfare participation. This estimate is large and significant until age 27 after which it converges to zero. The criminal record also has persistent effects for several adulthood outcomes. In particular, it increases the serious offending probability. This estimate is consistently positive and nearly always significant for the observation period. The criminal record further reduces the reliance on social welfare between ages 16 and 19 and increases the probability of drug use from age 22 onwards.

In Figure 4.5 we show the age-varying estimates for dynamic contagion in the adult outcomes for females. When compared to the males we find less variation during adulthood. An interesting finding is that intimate relationships reduce participation in social welfare programs during adolescence. This shows that social welfare rates for females can be reduced by both intimate relationships and, as shown in Figure 4.9, education. A second noticable finding is that drug use reduces the probability of intimate relationships from age 20 onwards, whereas social welfare increases the probability of intimate relationships until age 20. This again highlights the important and complicated role that relationships play in the lives of the females, much more so than for males.
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

4.5 Policy implications

While the parameter estimates in Section 4.4 provide information for the marginal importance of the various components of the model from Section 4.2, it is hard to see how the components simultaneously interact in their determination of the adult outcomes. In particular, the payoffs from investments in childhood skills are hard to quantify as complementing effects from signals and dynamic contagion in adult outcomes may increase or decrease subsequent payoffs. To be able to understand these dynamics we present some counterfactual analyses that quantify effects from shocks to the childhood skills on adult outcomes. We emphasize that the skills can have direct effects on the adult outcomes, via the improvement of the childhood signals and via multiplier effects from dynamic contagion during adulthood.

To study shocks to the cognitive and non-cognitive skills we proceed as follows. Given the estimated parameters we extract the posterior mean of the childhood skills
Figure 4.5: Age-varying estimates $A_t$ for females, which capture the dynamic contagion in the adult outcomes. The dotted lines indicate the 90% confidence regions.

$F_i$, for $i = 1, \ldots, N$. Given this estimate $\tilde{F}_i$, we compute baseline predictions for the outcome variables for each age period. In particular, we compute

$$\hat{Y}_i = \mathbb{E}(Y_i | \tilde{F}_i) = \int_{e_i} \mathbb{E}(Y_i | e_i; \tilde{F}_i) p(e_i) \, de_i,$$

where $\mathbb{E}(Y_i | e_i; \tilde{F}_i)$ follows from the mean of the binary distribution and the integral is computed by simulation methods. We emphasize that we regard the observed variables as a particular realization of the model and that we do not condition on the paths of these outcome variables. Such analyses are interesting if we would shock adult outcomes, via $e_{i,t}$, at some point during adulthood. Then the observations that were observed prior to shocking the model at time $t$ would be important.

Further, since we are not interested in predictions for a particular individual we average the predictions over the individuals. The resulting estimates are shown for
males in Figure 4.10 and for females in Figure 4.11 for all outcome variables. We also show the 90% confidence bounds for these model-implied rates. Next, we impose a positive shock to either the cognitive skills or the non-cognitive skills of one standard deviation. This increased skill vector is then inserted in (4.14) and the expectations for the outcome variables are recomputed. In Figures 4.10 and 4.11 we show the resulting average estimates for the adult outcomes for males and females.

We find that for males non-cognitive skills significantly reduce the probability of serious offending for the entire observational period. We notice that the reductions are much larger when compared to the marginal effects shown in Figure 4.6. A channel that explains the additional drop in the serious offending rate is the following. Non-cognitive skills significantly reduce the intensity of the criminal record (Table 4.5). The reduced criminal record then further lowers adulthood serious offending according to
Cognitive
\[ \text{(i)} \]
Non-Cognitive
\[ \text{(ii)} \]
\[ \text{(iii)} \]
\[ \text{(iv)} \]
\[ \text{(v)} \]

Figure 4.7: Age-varying estimates $B_t$ for females, which capture the persistent effects from the cognitive and non-cognitive skills. The dotted lines indicate the 90% confidence regions.

Figure 4.8. Many other paths for the effects from non-cognitive skills can be found by combining the marginal estimates from the previous sections. The main finding is that the combination of the marginal estimates implies that on average small improvements in non-cognitive skills can reduce serious offending by 4.6% per age-year between ages 17 and 32. Cognitive skills also reduce serious offending, but the estimates are only significant between ages 17-20 and 29-32.

Both shocks to cognitive and non-cognitive skills significantly increase the employment probability from age 23-24 onwards for males. On average per age-period the one standard deviation shocks imply gains of 4.1% and 5.4% for the cognitive and non-cognitive skills, respectively. We emphasize that these estimates are for the entire observational period. At age 32 the employment participation rate is increased on average by 8.3% (cognitive) and 8.7% (non-cognitive). For social welfare participation,
For females we find only significant results from the non-cognitive shock. In particular, until age 28 a one standard deviation positive shock to the non-cognitive skills significantly reduces the serious offending rate. The reductions are on average between 1.3% and 4.1%. The employment participation rate is significantly increased over the entire observational period. The average increase is 5.2% per age-year between ages 17 and 32. Finally, a positive shock to the non-cognitive skills significantly reduces the intimate relationships rate until age 21. Especially, for ages 16 and 17 higher non-cognitive skills imply large reductions in the relationship rate (13.4% and 13.9% respectively).

Finally, we consider one additional policy experiment. Suppose that we could impose employment at age 22 for all individuals. More specifically, we assume that we can facilitate employment for more than three months for all individuals at age 22. With
Figure 4.9: Age-varying estimates $C_t$ for females, which capture the persistent effects from the childhood signals. The dotted lines indicate the 90% confidence regions.

We find that, consistent with the findings in Figures 4.4 and 4.5, the serious offending probabilities are initially reduced for ages 23 to 25. The initial drop at age 23 is also larger when compared to the drop created by the non-cognitive skills improvement (Figures 4.10 and 4.11). However, the reduction in offending is rather short-lived and by age 26 the predicted offending rates are back to the model implied levels.

Similar results are found for the predicted employment probabilities. After an initial increase in the employment probability between ages 23 and 25, the rates drop back to the model implied rates. This highlights the difficulties for disadvantaged youths to remain continuously employed. The results for the other adult outcomes, social welfare, drug use and intimate relationships, are not significant for both males
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

Figure 4.10: Simulation results from increasing cognitive, or non-cognitive, skills by one standard deviation for males. The baseline rate is based on the posterior mean for the childhood skills. The standard deviation increase in considered from this average.

i. serious offending, ii. regular employment, iii. social welfare, iv. drug use, v. intimate relationships.

and females. We may conclude that increasing non-cognitive skills prior to age 16 has larger effects for disadvantaged youths on multiple adult outcomes when compared to creating employment at age 22.

4.6 Conclusion

In this paper we have proposed a reduced form nonlinear dynamic panel data model that decomposes multiple socioeconomic adult outcomes into payoffs from childhood skills and signals, and dynamic contagion in adult outcomes. The model included a factor structure to estimate cognitive and non-cognitive skills, which we allowed to
Figure 4.11: Simulation results from increasing cognitive, or non-cognitive, skills by one standard deviation for females. The baseline rate is based on the posterior mean for the childhood skills. The standard deviation increase in considered from this average. i. serious offending, ii. regular employment, iii. social welfare, iv. drug use, v. intimate relationships.

affect the childhood signals as well as the adult outcomes. The childhood signals were also allowed to affect the adult outcomes according to an age-varying payoff vector. The model was estimated using novel Monte Carlo maximum likelihood methods which enabled the simultaneous estimation of all parameters that pertained to the childhood and adulthood models.

We illustrated the model for samples of disadvantaged youths, who had been institutionalized in their teenage years. For these youths we found significant lasting effects from childhood skills on predominantly serious offending and employment. These marginal payoffs were initially small, but their eventual payoff at the end of adulthood was large due to multiplier effects from childhood signals and adulthood life transitions. The results clearly illustrate the interaction between early investments and later life
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

Figure 4.12: Simulation results from imposing employment at age 22 for males. The baseline rate is based on the posterior mean for the childhood skills. i. serious offending, ii. regular employment, iii. social welfare, iv. drug use, v. intimate relationships.

outcomes for both males and females.
Figure 4.13: Simulation results from imposing employment at age 22 for females. The baseline rate is based on the posterior mean for the childhood skills. i. serious offending, ii. regular employment, iii. social welfare, iv. drug use, v. intimate relationships.
Appendix A: details for the parameter estimation method

In this appendix we provide the details for the parameter estimation method that is outlined in Section 4.2.4. In particular, we consider the estimation of the posterior modes $\hat{e}_i$ and $\hat{F}_i$ and detail the construction of the importance densities $g(e_i|Y_i, \hat{F}_i)$ and $g(F_i|Y_i; \hat{e}_i)$.

To maximize $\log p(F_i, e_i|Y_i)$ with respect to $F_i$ and $e_i$ we make the following observations. When $F_i$ is fixed at some value, the conditional posterior mode of $e_i$ can be located based solely on $Y_i$, the adulthood observations. When $e_i$ is fixed at some value, the conditional posterior mode of $F_i$ depends on $Z_{i,C}, Z_{i,N}, E_i, O_i$ and $Y_i$. The vectors $Z_{i,C}, Z_{i,N}$ and $Y_i$ have larger dimensions when compared to $F_i$. This suggests that the dimension reduction techniques outlined in Jungbacker and Koopman (2014) and Mesters and Koopman (2014) can be used to obtain low-dimensions vectors that summarize sufficient information for the computationally efficient integration of $F_i$.

In particular, let $Z_i = (Z_{i,C}, Z_{i,N})'$, $\mu = (\mu_C, \mu_N)'$ and $\epsilon_i = (\epsilon_{i,C}, \epsilon_{i,N})'$, the factor model for the skills can now be written as

$$Z_i = \mu + \Lambda F_i + \epsilon_i, \quad \epsilon_i \sim IID(0, \Omega), \quad (4.15)$$

where

$$\Lambda = \begin{bmatrix} \alpha_C & 0 \\ 0 & \alpha_N \end{bmatrix},$$

and $\Omega$ is the diagonal variance matrix. Since $Z_i$ is of length $K_C + K_N > 2$ we can construct a transformed lower-dimensional vector $Z_i^l$ such that $E(F_i|Z_i) = E(F_i|Z_i^l)$. A convenient transformation that preserves this property is given by

$$Z_i^l = C_z\Lambda^\prime \Omega^{-1} (Z_i - \mu), \quad C_z = \text{choleski} \left[(\Lambda^\prime \Omega^{-1} \Lambda)^{-1}\right], \quad (4.16)$$

where $C_z$ is lower triangular. The resulting model for $Z_i^l$ is given by

$$Z_i^l = C_z^{-1} F_i + u_i, \quad \epsilon_i^l \sim IID(0, I), \quad (4.17)$$

where $Z_i^l$ now has dimension 2. The proof for $E(F_i|Z_i) = E(F_i|Z_i^l)$ is given in Jungbacker

We proceed by making a similar transformation for the model implied by \( p(Y_{i,t}|F_i, e_i) \)

First, we linearize \( \log p(Y_{i,t}|F_i, e_i) \)

\[
Y_{i,t} = c_{i,t} + B_t F_i + e_{i,t} + u_{i,t}, \quad u_{i,t} \sim N(0, D_{i,t}), \tag{4.18}
\]

where \( D_{i,t} \) is diagonal by construction and the construction of the coefficients \( c_{i,t} \) and \( D_{i,t} \) is discussed below. The conditional density for model (4.18) is given by

\[
g(Y_{i,t}|F_i, e_{i,t}) \equiv NID(c_{i,t} + B_t F_i + e_{i,t}, D_{i,t}) \tag{4.19}
\]

Given the linearized model (4.18) and a particular value for \( e_i \), say \( e_i^* \), we can obtain a lower dimensional vector \( Y^l_i \) such that \( E(F_i|Y_i; e_i^*) = E(F_i|Y^l_i; e_i^*) \). Similar as in (4.16) we define

\[
Y^l_i = C_y B' D^{-1}_{i} (Y_i - c_i - e_i^*), \quad C_y = \text{choleski} \left[(B'D^{-1}_i B)^{-1}\right], \tag{4.20}
\]

where \( B = (B'_1, \ldots, B'_T)' \), \( D_i = \text{diag}(D_{i,1}, \ldots, D_{i,T}) \) and \( c_i = (c'_{i,1}, \ldots, c'_{i,T})' \). The resulting model for \( Y^l_i \) is given by

\[
Y^l_i = C_y^{-1} F_i + \xi^l_i, \quad \xi_i \sim N(0, I), \tag{4.21}
\]

where the dimension of \( Y^l_i \) is \( M \times 1 \), \( M \) being the number of adult outcomes considered.

Finally, we still need to obtain a linear model for the childhood signals \( G_i = (E_i, O_i)' \). The linearized models for \( p(E_i|F_i) \) and \( p(O_i|F_i) \) are given by

\[
G_i = c_{G,i} + F_i + u_{G,i}, \quad u_{G,i} \sim N(0, D_{G,i}), \tag{4.22}
\]

where the construction of the coefficients \( c_{G,i} \) and \( D_{G,i} \) follows below. The density for the approximating model (4.22) is given by

\[
h(G_i|F_i) \equiv N(c_{G,i} + F_i, D_{G,i}). \tag{4.23}
\]
CHAPTER 4: CHILDHOOD SKILLS, SIGNALS AND SOCIOECONOMIC ADULTHOOD OUTCOMES

Summarizing, given \( e_i = e_i^* \) we can base our inference for \( E(F_i|Y_i; e_i^*) \) on

\[
\mathcal{Y}_i = \begin{bmatrix} C_z^{-1} \\ C_y^{-1} \\ I_2 \end{bmatrix} F_i + \xi_i, \quad \xi_i \sim N(0, \Gamma_i),
\]

(4.24)

where \( \mathcal{Y}_i = (Z_i^t, Y_i^t, G_i - c_{G,i})' \) and \( \Gamma_i = \text{diag}(I_M, I_2, D_{G,i}) \).

With these approximating models in place, we are now ready to detail the computation of the posterior modes. This is done using the following Gauss-Newton algorithm.

**Algorithm A**

1. Initialize the algorithm by choosing \( F_i^* \) and \( e_i^* \) as starting values;

2. Given the set of two equations

\[
\frac{\partial \log p(Y_{i,t}|F_i, e_{i,t})}{\partial (B_t F_i + e_{i,t})} = \frac{\partial \log g(Y_{i,t}|F_i, e_{i,t})}{\partial (B_t F_i + e_{i,t})},
\]

\[
\frac{\partial^2 \log p(Y_{i,t}|F_i, e_{i,t})}{\partial (B_t F_i + e_{i,t}) \partial (B_t F_i + e_{i,t})'} = \frac{\partial^2 \log g(Y_{i,t}|F_i, e_{i,t})}{\partial (B_t F_i + e_{i,t}) \partial (B_t F_i + e_{i,t})'},
\]

for \( i = 1, \ldots, N \) and \( t = 1, \ldots, T \), where \( p(Y_{i,t}|F_i, e_{i,t}) \) is the observation model and \( g(Y_{i,t}|F_i, e_{i,t}) \) is given by (4.19), we can deduce expressions for \( c_{i,t} \) and \( D_{i,t} \) as functions of \( F_i \) and \( e_{i,t} \), and compute \( c_{i,t} = c_{i,t}^* \) and \( D_{i,t} = D_{i,t}^* \) for \( e_{i,t} = e_{i,t}^* \) and \( F_i = F_i^* \);

3. Given \( c_{i,t} \) and \( D_{i,t} \) we can compute the lower dimensional model (4.21);

4. Next, given the set of two equations

\[
\frac{\partial \log p(G_i|F_i)}{\partial F_i} = \frac{\partial \log h(G_i|F_i)}{\partial F_i}, \quad \frac{\partial^2 \log p(G_i|F_i)}{\partial F_i \partial F_i'} = \frac{\partial^2 \log h(G_i|F_i)}{\partial F_i \partial F_i'},
\]

for \( i = 1, \ldots, N \), where \( p(G_i|F_i) \) is the observation model and \( h(G_i|F_i) \) is given by (4.23), we can deduce expressions for \( c_{G,i} \) and \( D_{G,i} \) as functions of \( F_i \), and compute \( c_{G,i} = c_{G,i}^* \) and \( D_{G,i} = D_{G,i}^* \) for \( F_i = F_i^* \);

5. Compute \( \hat{F}_i = E(F_i|\mathcal{Y}_i; e_i^*) \) from model (4.24);
6. Replace \( F_i^* \) by \( F_i^* = \tilde{F}_i \);

7. Compute \( \tilde{e}_i = \mathbb{E}(e_i|Y_i; F_i^*) \) from model (4.18)

8. Replace \( e_i^* \) by \( e_i^* = \tilde{e}_i \)

9. Iterate from (2) to (9) until convergence.

Since the mode and the mean of the approximating linear Gaussian model are set equal to the mode of the original model, it holds that \( \tilde{F}_i = \hat{F}_i = \arg\max_{\mu} p(F_i|Y_i; \hat{e}_i) \) and \( \tilde{e}_i = \hat{e}_i = \arg\max_{\mu} p(e_i|Y_i; \hat{F}_i) \). Further, it holds that \( \{\tilde{F}_i, \tilde{e}_i\} = \arg\max_{F_i, e_i} p(F_i, e_i|Y_i) \).

Now that we have obtained the posterior modes the importance densities follow from the approximating models (4.24) and (4.18). In particular, given \( \hat{e}_i \) and the final coefficients for this model from Algorithm A we can use multivariate regression methods to draw conditional samples from the model (4.24). Also, given \( \hat{F}_i \) we can use the simulation smoother methods of Durbin and Koopman (2002) to sample from the model (4.18). More details for the outlined estimation method can be found in Mesters and Koopman (2014).

**Appendix B: construction of the cubic spline functions**

In this appendix we provide the details for the construction of the cubic splines that we use to model the age-varying payoff matrices \( A_t, B_t, C_t \) and \( D_t \). More details for methods using splines can be found in Poirier (1976). In principal, it is possible to treat all the individual parameters in \( A_t, B_t, C_t \) and \( D_t \) as deterministic parameters and estimate them along with the other parameters. However, since the time series dimension is \( T = 16 \) this would imply \( 16 \times (25 + 10 + 10) = 720 \) deterministic model parameters only from the payoff matrices. Optimizing the Monte Carlo likelihood using numerical methods over such a large parameter space is practically not feasible.

To avoid this problem, we make the assumption that the payoffs vary smoothly with age. This allows us to fit cubic splines between the subsequent elements of the payoff matrices. These rely on a smaller number of parameters. In particular, we seek a subset of \( K < T \) knots from the set \( \{1, \ldots, T\} \) between which we fit cubic spline
functions. The number of knots and the location of the knots can be determined in a variety of ways (Jungbacker et al., 2014). In this paper we set the locations equal to ages 16, 24 and 32. Experiments with more knots have led to similar results.
Chapter 5

The Effect of Unemployment on Crime in High Risk Families in the Netherlands between 1920 and 2005

5.1 Introduction

The relationship between the labor market and criminal behavior has been of long-standing interest in the social sciences. The economic analysis of this relationship started with the seminal contributions of Becker (1968) and Ehrlich (1973). The intuitive appeal of the argument that improving labor market conditions cause individuals to commit less crime is apparently so self-evident that the empirical evidence should be overwhelming. However, only in roughly the last decade have researchers been able to document reliable significant effects of various labor market conditions on crime rates. By employing novel panel data strategies Doyle, Ehsan Ahmed, and Horn (1999), Raphael and Winter-Ebmer (2001), Gould et al. (2002), Papps and Winkelmann (2002), Machin and Meghir (2004), Ihlanfeldt (2007) and Lin (2008) have been able to identify causal effects of labor market outcomes on crime rates. A typical analysis by one of the aforementioned authors exploits geographical variation in crime rates and labor market opportunities to estimate the effect of the labor market outcomes, while controlling for a host of confounding variables. Further, to establish the direction
of causality\textsuperscript{1} various combinations of instrumental variables have been suggested to instrument for the labor market outcomes. Mustard (2010) surveys the economic-crime literature, with an emphasis on modeling difficulties and international comparison. 

To this date, little evidence has been produced for a relationship between crime and labor market opportunities in the Netherlands. Jongman (1982), Jongman (1988) and Ploeg (1991) discuss this relationship from various angles, but their work is mainly descriptive. Further, using individual-level data, Miedema (1997), van der Geest et al. (2011) and van der Geest (2011) investigate the criminal and employment careers of high risk males. From their work it can be concluded that employment is likely to be negatively correlated with criminal behavior for low-skilled males. However, the direction of causality that is implied remains unclear.

This paper examines the degree to which changes in the conviction rates of 181 high risk families from the Netherlands can be explained by changes in the national unemployment rate. The families consist of four generations, that are observed between 1920 and 2005. The first generation consists of males who were institutionalized in a reform school in their teenage years. They come from low social classes and they and their descendants are at high risk of offending compared to the average Dutch population. Bijleveld et al. (2007) discuss the sample and its origins in detail. Their characteristics makes these families a very interesting subgroup of the population to examine, as Grogger (1998), Gould et al. (2002) and Machin and Meghir (2004) find that unskilled workers are more likely to commit crimes when their labor market prospects decrease. In this paper we establish whether the national unemployment rate has a causal effect on the criminal behavior of the families and whether this effect varies over time.

Very few studies so far have examined crime and unemployment trends over such a long time span - 85 years. This approach gives a chance to challenge the perception that the association between crime and unemployment may be unchanging. It is rather likely that the relationship is influenced by for example societal and political changes. Furthermore, the family-level of analysis provides several modeling benefits compared to the standard individual-level or macro-level analyses. First, it allows us to control for demographic and socio-economic status variables at the family level.

\textsuperscript{1}This is necessary as the direction of causality in the relationship between labor market outcomes and crime rates is not known a priori. Crime rates can causally influence labor market outcomes and vice versa, see Grogger (1998) and Raphael and Winter-Ebner (2001) for a more elaborate discussion.
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

Second, the constructed family-specific time series are more informative compared to individual-specific time series. Third, the family-level avoids the explicit modeling of correlations between family members, which is shown to be quite complicated even in a static framework, see Durlauf and Ioannides (2009). Fourth, simultaneity problems, thus reverse causality from crime on unemployment rate, are virtually non-existent as the effect of the criminal behavior of the families on the national unemployment rate is negligible. Fifth, we limit a bias caused by omitting of variables by controlling extensively for national trends in criminal opportunities, justice policies and criminogenic factors, as a large sample of these measures is available at the national level.

The analyses in this paper are divided in two parts. First, we estimate a large variety of model specifications to establish the overall robustness of the effect of the unemployment rate on the conviction outcomes. This is done for two time periods: 1920-2005 and 1950-2005. We consider three samples of convictions. A full sample containing all convictions for serious crimes and two samples containing only property or violent convictions. The models that we consider are included in the class of generalized dynamic panel data models proposed by (Mesters & Koopman, 2014). The estimated models differ with respect to the inclusion of control variables. Three types of control variables can be distinguished in our models. First, a set of family-specific control variables is included to control for the demographic and economic composition of the family. Second, a set of macro-level trends is included to control for pro-cyclically varying factors. These factors have been categorized by Cook and Zarkin (1985) and their exclusion can understate or overstate the effect of the unemployment rate, see the discussion in Raphael and Winter-Ebmer (2001). Third, we exploit the features of our panel data by including time-invariant family-specific effects, time-varying common effects and lagged offending outcomes. The latter capture the causal effect of crimes committed in the family in the previous time period, see also Machin and Meghir (2004).

In our second analysis we seek to establish whether similarities in the fluctuations in the conviction rates and the national unemployment rate can be considered due to a changing causal effect. The two rates show more similarities after 1950, see Figure 5.3. To formally investigate whether a changing effect is present, we estimate a time-varying regression effect for the unemployment rate, see Durbin and Koopman (2001, Chapter 3). Between 1920 and 2005 Dutch society changed in many respects. For
example, society has evolved and consumerism rather than frugality has become the norm, implying that the possession and flaunting of material goods has become more important than modesty and restraint. Also, social control through institutions, such as the church and the extended family or the neighborhood has strongly declined. The impact of these secular changes on the effect of unemployment is unknown. We discuss what factors might be causing the changes.

The remainder of this paper is organized as follows. The next section describes the data. Special interest is in describing the trends that exist in the conviction outcomes and the unemployment rate. Section 5.3 discusses our main empirical strategy. Here we describe our basic panel data models and discuss our control variables. In Section 5.4 we allow for time-varying effects of the unemployment rate. In Section 5.5 we present our conclusions and possibilities for further research.

5.2 Data

The crime data originates from the TRANS-5 study, see Bijleveld et al. (2007). The TRANS-5 dataset consists of observations for five generations of families, G1-G5, as shown in Figure 5.1. Permission for the study was obtained from the legal successor of the Harreveld institution, the Frentrop foundation, as well The Netherlands Minister of Justice. The five generations are ancestors and descendants of the original sample (G2), which consists of 198 males who were institutionalized in a Catholic reform school between 1911 and 1914. They were sent there for various reasons, such as criminal behavior or disrupted family situations. In the reform school, they received, if necessary, elementary education and stayed until they had finished accredited vocational training to become, for example, a carpenter, baker, painter, or shoemaker. The offspring of the G2 males and their spouses were traced in Dutch genealogical and municipal records, resulting in a 100% retrieval rate. Sample members of G2 who had emigrated, or died before the age of 21, were considered lost to follow-up and their descendants were not traced. After removing these, 181 men remained who had offspring that is labeled G3; subsequent generations are labeled as G4 and G5.

In this paper, we rely on four generations, G2-G5, as criminal data for the G1 generation was available but was probably incomplete. In total, we included 4,120 individuals, to which an additional 1,919 spouses could be linked, resulting in 6,039
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

Figure 5.1: Design of the five-generation sample

men and women. Each individual is included from age 12 to 60. The lower limit is chosen as 12 years is the age of legal accountability in the Netherlands. The upper bound is chosen for two reasons. First, little criminal acts can be expected from older individuals (the records also show this) and second, until recently, most individuals quit working around age 60, either voluntary or forced. Observations between 1920 and 2005 are included for generations G2 to G5. The individual-level information on these individuals is aggregated to extended family level, where the 181 males are considered as the original family members. The 181 male members were on average 21 years of age in 1920. The family data and its retrieval procedure are described below.

5.2.1 Conviction data

For all sample members born after 1916, judicial information was collected from the archives of the Dutch Criminal Records Documentation Service. As a result, detailed information concerning the complete criminal career for the generations G3, G4 and G5 is available. For those born before 1916 (G2), information about delinquency was gathered from several sources: district court archives, police archives, prison archives and beggars’ colonies’ records. We will consider three different samples of convictions: serious convictions, property and violent convictions. We operationalized serious crim-
inal behavior according to the definition of Loeber et al. (1998). Sample members who had been convicted for one or more of the following crimes were defined as serious offenders: violent offenses, property offenses, drugs offenses, arson and violations of weapons and firearms regulations. In constructing our dependent variable, only those registrations were used that were not acquitted or dismissed for ‘technical’ reasons (predominantly because the prosecutor deemed that insufficient proof was available). Consequently, the collected data represent the lower limit of actual offending. We time delinquency to the date the offense was committed. If no date was known, we estimated it as date of conviction minus the average duration of disposition, set at one year. If no disposition date was known we estimated it at 1 July of the year of registration. Following these rules, we therefore always classify crimes in the manner in which the last criminal justice institution that dealt with the case classified it. More information may be found in Bijleveld et al. (2007). Prevailing definitions of the period under investigation were used to define acts as delinquent, see Bijleveld et al. (2007). As data on matched controls are also available, we are able to assess that these respondents are indeed high-risk: around 50 per cent of the G3, G4 and G5 men were convicted for at least one offense against around 20 per cent of control men, the women are also at elevated risk, see Bijleveld et al. (2007).

In Figure 5.2 we show the sample conviction rate (left column). This is the number of individuals who got convicted for a serious, property or violent crime in each year, divided by the total number of individuals present in each year. For example in 2005, approximately 2% of the individuals present in the sample was convicted for a serious crime. In the period between 1920 and 1930 higher levels of convictions are found. This is caused by the fact that fewer individuals are present in these periods (approximately 400-500) and that the majority of these are young males. The serious convictions rate is similar to the property convictions rate as a large proportion of the serious convictions are property convictions. In later years this close connection disappears, although the trends remain similar. The violent sample conviction rate fluctuated more highly, but the rate is generally low.

In the right column of 5.2 we show the mean residual conviction rates and the smoothed levels, for serious, property and violent convictions. The residuals are obtained after running family-level OLS regressions, correcting for the demographic (age, gender and family) and economic (socioeconomic status) composition of the families.
Figure 5.2: The left column shows the sample conviction rate for serious, property and violent convictions. The rate is computed as the number of individuals convicted for serious, property, or violent crimes, divided by the total number of individuals between 12-60 years of age in the corresponding year. The right column shows the residual sample conviction rate. The residuals were computed by running OLS regressions on the family level. The dependent variables were the sample conviction rates per family. The regressions included a constant, four age variables (the ratio of 12-17, 18-24, 25-34 and 35-44 year olds), the ratio of males, the ratio of family relationships (father-mother, father-son, father daughter) and the socio-economic status indicators.

The exact construction of these variables is discussed below. The smoothed level is computed by applying the Kalman filter and smoothing recursions to the residuals using a simple local level model, see Durbin and Koopman (2001). The smoothed level can be seen as the mean level of the convictions. The heavy fluctuations in the ‘raw’ conviction data are now almost completely removed.
5.2.2 Additional family information

Next to the conviction data, we have some demographic and socioeconomic status information at our disposal. The demographic information was traced in the Dutch municipal and administrative records. First the members of the original sample (G2) were identified and their ‘gezinskaarten’ (family cards) traced a family-based registration system operational from the first decades of the 20th century until just before the Second World War. In this system all family members residing at one address were registered on one card. This gave us the family composition, and information on for example the professions of the head of household.

The information on the professions was supplemented by screening the non-medical part of the Netherlands Ministry of Defense DARIC archives. In principle all men from birth year 1930 onwards until 1996 in the Netherlands were screened for military service; the screening contained a ‘family interview’ in which the father’s profession was listed. The screening records of any descendants of the 181 men thus contain information on the profession of their father. Permission was obtained from the Netherlands Ministry of Defense for this part of the data collection.

From the municipal archives and the non-medical part of the military records we constructed family level control variables. To control for the changing age composition within the families control variables were constructed as the ratio of family members who were between the ages (in years) 12-17, 18-24, 25-34 and 35-44. The final age group ratio, 45-60, was left out to avoid multicollinearity. A control for the gender composition of the family was computed as the ratio of males in each family in each year. Further we aim to control for the family relations within each family. We constructed variables for the number of father-mother, father-son, father-daughter, mother-son and mother-daughter relationships that were found in each family for each year. The numbers were scaled with respect to the total number of family members. Finally, a socioeconomic status variable was constructed by recoding the professions of the male members. In particular, all occupations were coded into occupational class categories according to the Historical International Standard Classification of Occupations (HISCO) classification, see Van Leeuwen, Maas, and Miles (2002). This was done as our observation period is 85 years containing thus contemporary and historical occupations. HISCO offers the chance to standardize the changes in occupations over time.
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

and make occupational statuses comparable. HISCO occupational codes were then grouped into seven social class categories (HISCLASS): 1 lower-skilled and unskilled farm workers, 2 unskilled workers, 3 lower-skilled workers, 4 farmers and fishermen, 5 foremen and skilled workers, 6 lower managers and professionals, clerical and sales personnel, 7 higher managers and professionals. Men in our data were assigned the highest social class occupation we had on them (either from DARIC or gezinskaart information). For the females in our data no information on occupations is available. We assigned them the socioeconomic status of their father and when they married the socioeconomic status of their husband. This seems an appropriate method for earlier years when females did not work very often, but much less so for recent years.

5.2.3 Relation to the national unemployment rate

In Figure 5.3 we show the mean corrected unemployment rate together with the filtered conviction signals, which are the same as in Figure 5.2. The unemployment rate is obtained from Statistics Netherlands and contains information on registered unemployment. The unemployment rate shows large fluctuations between 1920 and 2005. Two periods of elevated unemployment can be distinguished. First, in the 1930’s during the Great Depression the unemployment rate increased to nearly 20%. In this period the conviction rates remained largely unchanged. This holds for serious, property and violent convictions. The rather remarkable none-response of the conviction rates during the great depression in also found at the national level. Leistra and Nieuwbeerta (2003) confirm this for homicide rates, which are probably the best recorded historical time series, but are likely to be the least responsive to changes in the unemployment rate. Jongman and Henkes (1981) find little responses for a broader sample of crime types during the great depression. They concluded that the increased levels of poverty did not correspond to the unchanged crime rates. The consistency of these findings leads us to believe that we are not looking at an artifact of our own data sample.

Internationally, similar results for the great depression period have been documented. Fishback, Johnson, and Kantor (2010) find no increases in property crime rates for a sample of 114 US cities. They argue that this is caused by large increases in government spending in this period. In The Netherlands no such increases in government spending are documented, see den Bakker (2008). The national rate of offenses
known to the police in England and Whales is also unaltered during the 1930’s, see Hicks and Allen (1999). Similar findings are documented for Belgium in (Francois, 2008).

The conviction rate remained apparently stable even as the unemployment rate fall again from its peak. The unemployment rate declined to around 3% for most of the 1950’s, 1960’s and 1970’s.

The second period of increasing unemployment in The Netherlands takes place in the 1980’s. In this period the unemployment rate increases to approximately 10%. Interestingly, all conviction rates of the 181 families seem to show a response to this increase, although this holds more for property crimes (theft, burglary, embezzlement fraud) and much less so for violent crimes (mainly: assault, sex offenses, extortion, threat and homicide). In international studies similar findings have been documented for this recent increase, see Doyle et al. (1999), Raphael and Winter-Ebmer (2001) and Papps and Winkelmann (2002). To sum, prior to 1950 little relation can be observed between the conviction rates and the unemployment rate. Conviction rates remain largely stable while the unemployment rate first increases rapidly after which it decreases. The overall pattern that emerges leads us to suspect that the relationship between (property) crime and the unemployment rate is non-existent, or weak, prior to 1950 but emerges after this period. However, the similarities between the signal for property crime and the unemployment rate after 1950 are not necessarily the result of any causal connection. The next sections seek to determine whether a causal effect of the unemployment rate on the conviction outcomes exists.
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

Figure 5.3: From top to bottom we show the serious, property and violent conviction signals, as computed in Figure 5.2, together with the mean corrected national unemployment rate, which is obtained from Statistics Netherlands (www.cbs.nl).

5.3 Empirical panel data strategy

In this section we discuss our basic panel data models that we used to identify the effect of the unemployment rate on the conviction outcomes of the families. The models included fall in the class of generalized dynamic panel data models proposed by Mesters and Koopman (2014). We propose a large variety of models, which differ with respect to the inclusion of control variables. Further, we estimate the models for three different groups of convictions; serious, property and violent, and two different time periods; 1920-2005 and 1950-2005. In this manner we aim to establish the robustness of the overall effect of the unemployment rate on the convictions.
5.3.1 The observation model

There are \( N = 181 \) families, with each family indexed by \( i \) for \( i = 1, \ldots, N \). In each year \( \tau_t \) there are \( n_{i,\tau_t} \) members in family \( i \), for \( t = 1, \ldots, T \), with \( T = 85 \), where \( \tau_1 = 1921 \) and \( \tau_{85} = 2005 \). The number of individuals who are convicted from family \( i \) in year \( \tau_t \) is denoted by \( y_{i,\tau_t} \). This can be seen as the number of ‘successes’ stemming from \( n_{i,\tau_t} \) trials. Therefore, the observations \( y_{i,\tau_t} \), for \( i = 1, \ldots, N \) and \( t = 1, \ldots, T \), are modeled by the binomial density given by,

\[
y_{i,\tau_t} \sim \text{Binomial}(n_{i,\tau_t}, \pi_{i,\tau_t}), \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{5.1}
\]

where \( n_{i,\tau_t} \) is the number of family members and \( \pi_{i,\tau_t} \) is the conviction probability for an individual from family \( i \) in year \( \tau_t \). Of main interest in this paper is the effect that the unemployment rate has on the conviction probability. The conviction probability is restricted between zero and one. To avoid difficulties during estimation procedures we model the transformed conviction probability, \( \theta_{i,\tau_t} = \log\left[ \pi_{i,\tau_t} / (1 - \pi_{i,\tau_t}) \right] \), which is equivalent to the log odds ratio. We refer to \( \theta_{i,\tau_t} \) as the signal.

The conditional density for \( y_{i,\tau_t} \), given \( \theta_{i,\tau_t} \), is given by

\[
p(y_{i,\tau_t} | \theta_{i,\tau_t}, \psi) \equiv \exp \left[ y_{i,\tau_t} \theta_{i,\tau_t} - n_{i,\tau_t} \log(1 + \exp \theta_{i,\tau_t}) + \log \left( \frac{n_{i,\tau_t}}{y_{i,\tau_t}} \right) \right], \tag{5.2}
\]

where \( \psi \) is the parameter vector. Density (5.2) is discussed more elaborately in Durbin and Koopman (2001, Section 10.3.3). As \( n_{i,\tau_t} \) is known and fixed, density (5.2) is entirely determined by signal \( \theta_{i,\tau_t} \), which may depend on parameters \( \psi \). Further, density (5.2) is considered independent given signal \( \theta_{i,\tau_t} \), for all \( i = 1, \ldots, N \) and \( t = 1, \ldots, T \). All dynamics and variables are modeled through signal \( \theta_{i,\tau_t} \). It follows that

\[
p(y | \theta, \psi) = \prod_{i=1}^{N} \prod_{t=1}^{T} p(y_{i,\tau_t} | \theta_{i,\tau_t}, \psi), \tag{5.3}
\]

where \( y = \{ y_{i,\tau_t} \}_{i=1,\ldots,N, \ t=1,\ldots,T} \) and \( \theta = \{ \theta_{i,\tau_t} \}_{i=1,\ldots,N, \ t=1,\ldots,T} \).
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

5.3.2 The signal

The signal $\theta_{i,\tau_t}$ is of central interest in our model. In our base model for $\theta_{i,\tau_t}$ we assume that the unemployment rate affects $\theta_{i,\tau_t}$ similarly for each family $i$ and year $\tau_t$. We vary the construction of the signal with respect to the inclusion of control variables. The most general signal that we include is given by

$$\theta_{i,\tau_t} = \delta \text{UN}_{\tau_t} + x_{i,\tau_t}\beta + w_{\tau_t}\lambda + \gamma y_{i,\tau_{t-1}} + \mu_i + \xi_{\tau_t}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \quad (5.4)$$

where $\text{UN}_{\tau_t}$ is the unemployment rate in year $\tau_t$, $x_{i,\tau_t}$ is a vector of family-specific control variables, $w_{\tau_t}$ is a vector of time-varying common control variables, $y_{i,\tau_{t-1}}$ is the outcome from the previous time period, $\mu_i$ is the family-specific effect and $\xi_{\tau_t}$ is the time-varying effect.

The vector of family-specific control variables $x_{i,\tau_t}$ includes the four age variables, the ratio of males, the family relationship ratios, and the socioeconomic status indicator, see the discussion in Section 5.2. The age variables are constructed to control for changes in the age composition of the families. We include the ratios of the age groups (in years); 12-17, 18-24, 25-34 and 35-44, in each family in each year. The family relationship ratios are included to control for the fact that the within family conviction probabilities might not be independent. More specific, if the members of the families had independent and identical probabilities of getting convicted, density (5.1) would be entirely correct. This is however not the case. For example, if in a certain period two persons from one family are convicted of a crime it could be that the two outcomes had a causal affect on each other. To capture the impact of these within-family effects we include controls that account for the number of father-mother, father-son, father-daughter, mother-son and mother-daughter relationships. The socioeconomic status variable is averaged from the individual level socioeconomic status variables as discussed in Section 5.2. It is possible that the socioeconomic status variables are simultaneously determined with the conviction outcomes; therefore the corresponding coefficients may be biased. To investigate this we estimated models without these controls and also while including the first lag of the socioeconomic status variables. In all cases the estimates for the unemployment rate remained similar.

The vector of common control variables includes variables that are known to vary pro-cyclically with the unemployment rate and can possibly also affect the conviction...
outcomes. Excluding these may lead us to not actually measure the impact of unemployment, but rather some composition of similar varying national trends, see the discussion in Raphael and Winter-Ebmer (2001). Cook and Zarkin (1985) discuss four broad categories of such pro-cyclical aspects that are known to vary with the business cycle. They are: (1) variation in legal employment opportunities, (2) variation in criminal opportunities, (3) consumption of criminogenic commodities and (4) variation in the response of the criminal justice system. While the variation in legal employment opportunities is of central interest in this paper we aim to reduce the influence of the other categories. We include the national hourly wage and the logarithm of the gross domestic product to control for the general economic conditions in the Netherlands. Criminogenic commodities are items such as drugs, guns and alcohol. To control for influences of these we include beer consumption in liters. Unfortunately reliable drugs consumption rates are not available to us for our sample period. Given the unique position of drugs in the Netherlands its exclusion can be important. Guns have less relevance to the setting in The Netherlands. The variation in the response of the criminal justice system is measured by including the national conviction rate and the incarceration rate. By including the national conviction rate we make sure that the observed variation in the number of convicted family members is not due to nationwide changes in the criminal justice system, but rather due to changes in criminal behavior.

The family-specific effect $\mu_i$ is included to capture all time-invariant unobserved differences that exist between the families. The family-specific effect is given by

$$\mu_i \sim NID(\mu_0^\iota, \sigma_\mu^2), \quad i = 1, \ldots, N,$$

where $\mu_0^\iota$ is the overall mean and $\sigma_\mu^2$ is the variance. The normality assumption is standard in random effects models, see the discussion in Baltagi (2005). To make sure that the family-specific means $\mu_i$ can be correctly identified we standardize all family-specific and common variables, $x_{i,t}$ and $w_t$, to have mean zero and unit variance. The family-specific means are then interpretable as the mean family conviction log odds ratio with average covariates. Similar strategies to avoid biases from correlation between $x_{i,\tau_t}$, $w_{\tau_t}$ and $\mu_i$ are considered by Juarez and Steel (2010) and Mesters and Koopman (2014).

In addition to the family-specific effect we also include time-varying effect $\xi_{\tau_t}$. The
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

time-varying effect captures unobserved factors that affect the offending outcomes that are common to all families, but may vary over time. We model the time-varying effect by a random walk process, given by

$$\xi_{\tau t} = \xi_{\tau t-1} + \eta_{\tau t}, \quad \eta_{\tau t} \sim NID(0, \sigma^2_\eta), \quad t = 2, \ldots, T,$$

where the initial time-varying effect, $\xi_{\tau 1}$, is fixed to zero for identification purposes. Autoregressive processes were also investigated but did not lead to substantial improvements. The family-specific and time-varying random effects are considered independent for all $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The binomial dynamic panel data model is summarized by equations (5.2), (5.4), (5.5) and (5.6). The parameters are summarized in the parameter vector $\psi = \{\delta, \gamma, \beta, \lambda, \mu_0, \sigma_\mu, \sigma_\eta\}$.

5.3.3 Results

The parameters of the binomial panel data model are estimated for three types of convictions: serious, property and violent convictions, and for two time periods; the full time period 1921-2005 and the subsample 1951-2005. The data and trends were described in Section 5.2. Variables $y_{i,\tau 0}$ are fixed at their 1920 and 1950 values for $i = 1, \ldots, N$, respectively. The parameters are estimated using the Monte Carlo maximum likelihood methods developed in Mesters and Koopman (2014). The implementation of this method for the binomial dynamic panel data model is discussed in Appendix A.

Table 1.a presents the results for the models where the dependent variable, $y_{i,\tau t}$, is the number of members from each family convicted for serious offenses. The first three columns give the results for the period 1920-2005, while the next three give the results for the period 1950-2005. For each period we estimated three different models that vary with respect to the inclusion of control variables. They are summarized as; (a) including only statistical controls, (b) including statistical and family controls and (c) including statistical, family and macro-level controls. The results indicate that the effect of the unemployment rate is weak and becomes insignificant when controlling for macro-level trends. The effects found for the period 1950-2005 are stronger than for the 1920-2005 period.

Table 2.a presents the results for property convictions. The effect of unemployment
is positive and significant at the 5% level of confidence for all model specifications for the 1950-2005 sample period. The magnitude of the relationship indicates that a 1 percentage point decrease in the unemployment rate causes a decline in the property crime probability of between 2.9 and 4.6 percentage points. A decline of 3.8 percentage points is recorded for our model including all control variables. The magnitude of the found effect is comparable to the results found for property crime in (Lin, 2008) and somewhat higher compared to the results of Raphael and Winter-Ebmer (2001). For the period 1920-2005 the effect sizes are smaller and become insignificant when the macro-level control variables are added to the model.

The results for violent convictions are presented in Table 3.a. In the first specification, only including statistical controls, the coefficient for unemployment is positive and significant. However, when adding control variables to correct for family composition, the effect of the unemployment rate is not significant. This is found for both sample periods and remains so when including macro-level control variables. The effect of the unemployment rate even becomes slightly negative (insignificant) for the 1920-2005 sample period when the latter control variables are added. These results imply that the unemployment rate has no influence on violent offending. This is also found for US state level panel data by Raphael and Winter-Ebmer (2001) and (Lin, 2008).

Concerning the performance of the other variables listed in Tables 1.a,b,c, the family control variables seem important. Consistent with previous research on the age-crime profile, the conviction probabilities increase when more individuals below the age of 25 are present in the family. Also, higher ratios of males in the family increases conviction probabilities. This corresponds to the fact that the majority of crimes is committed by males. For violent crimes the documented age-crime profile is somewhat different. Increasing ratios of the youngest age group, consisting of 12-17 year olds, is found to decrease the conviction probabilities. This corresponds to the general finding that violent crimes are more likely to be committed by older males. For violent crimes the magnitude of the effect of the male ratio is found much higher compared to the effects found for serious and property crimes. The socio-economic status variable lowers the conviction probabilities for all crime categories. The family relationship variables indicate that increasing numbers of father-son and mother-son pairs in the family increases the conviction probability. Also increasing numbers of father-mother relationships decreases the conviction probability. This is consistent with criminological
life course theories that suggest that marriage lowers criminal propensity.

The controls for the common national trends are interesting in their own right. Much to our own surprise we are unable to identify significant effects of any other trend consistently. The estimated standard errors are much higher compared to those estimated for the unemployment rate. The estimates for the national conviction rate and the prison population are positive indicating their relationship to our dependent variables, which were measures of convictions. The effect for beer consumption is also positive. It is significant for serious and property crimes for the sample period 1920-2005. The control variables for general economic conditions, i.e. log GDP and log wages, are found to be insignificant in all models. Contrary to the findings of Grogger (1998), Gould et al. (2002) and Machin and Meghir (2004), the national wage rate does not seem an important determinant in the crime rates. This is possibly caused by the specific situation in the Netherlands, where minimum wages are typically high. This would imply that the mere fact of having a job would be more important compared to the actual wage that is received.

The statistical controls reveal some interesting information. The differences between the families are large for all three types of convictions. These differences remain visible for all models, despite the inclusion of the control variables. This is mainly caused by the fact that we standardized all variables to have mean zero and unit variance. The time-varying effects show low variance, which rapidly decreases when control variables are added. The state dependence variables are positive and significant in all model specifications. This indicates that some causal effects of previous crimes committed within the family exist. A more specific explanation for the meaning of this variable can possibly be found by analyzing conviction rates at the individual level. In this manner it would be possible to analyze which family members influence each other, or whether the same individuals are responsible for sequences of convictions.

Overall the results lead us to conclude that the effect of the unemployment rate is significant for property crimes for the period 1950-2005. This is the only sequence of models that has given us these consistent results. The magnitude of a 2.9% and 3.8% increase is found similar to (Lin, 2008). For all other crime types the effect of the unemployment rate is found lower in magnitude and often not significant.
### Table 1.a: Baseline parameter estimates for serious convictions

<table>
<thead>
<tr>
<th></th>
<th>Serious Convictions</th>
<th></th>
<th>1950-2005</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unemployment</strong></td>
<td>0.0180 0.0078</td>
<td>0.0216 0.0079</td>
<td>0.0001 0.0105</td>
<td>0.0350 0.0090 0.0439 0.0094</td>
</tr>
<tr>
<td><strong>Controls for trend</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-</td>
<td>-</td>
<td>-0.1329 0.5305</td>
<td>-</td>
</tr>
<tr>
<td>Inflation</td>
<td>-</td>
<td>-</td>
<td>-0.0109 0.1446</td>
<td>-</td>
</tr>
<tr>
<td>Wages</td>
<td>-</td>
<td>-</td>
<td>0.2047 1.3002</td>
<td>-</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>-</td>
<td>-</td>
<td>0.3956 0.1786</td>
<td>-</td>
</tr>
<tr>
<td>Conviction rate</td>
<td>-</td>
<td>-</td>
<td>0.2015 0.3049</td>
<td>-</td>
</tr>
<tr>
<td>Prison population</td>
<td>-</td>
<td>-</td>
<td>0.2519 0.1454</td>
<td>-</td>
</tr>
<tr>
<td><strong>Controls for Family</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 12-17</td>
<td>-</td>
<td>0.6877 0.2408</td>
<td>1.6326 0.3279</td>
<td>-</td>
</tr>
<tr>
<td>Age 18-24</td>
<td>-</td>
<td>1.5210 0.2402</td>
<td>2.1965 0.3161</td>
<td>-</td>
</tr>
<tr>
<td>Age 25-34</td>
<td>-</td>
<td>0.9289 0.2282</td>
<td>1.4752 0.2966</td>
<td>-</td>
</tr>
<tr>
<td>Age 35-44</td>
<td>-</td>
<td>0.3726 0.2324</td>
<td>0.4925 0.2699</td>
<td>-</td>
</tr>
<tr>
<td>Ratio males</td>
<td>-</td>
<td>1.5185 0.3386</td>
<td>1.7148 0.3455</td>
<td>-</td>
</tr>
<tr>
<td>SES</td>
<td>-</td>
<td>0.0256 0.0151</td>
<td>0.0246 0.0154</td>
<td>-</td>
</tr>
<tr>
<td><strong>Statistical controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-4.7949 0.0921</td>
<td>-6.2482 0.5015</td>
<td>-7.0193 0.5420</td>
<td>-4.9098 0.0948 5.9149 0.6359</td>
</tr>
<tr>
<td>State dependence</td>
<td>0.1501 0.0160</td>
<td>0.1679 0.0163</td>
<td>0.1486 0.0165</td>
<td>0.1315 0.0169 0.1391 0.0170</td>
</tr>
<tr>
<td>Family-specific</td>
<td>0.9556 0.0788</td>
<td>0.9305 0.0743</td>
<td>0.9387 0.0759</td>
<td>0.9328 0.0764 0.8940 0.0742</td>
</tr>
<tr>
<td>Time-varying</td>
<td>0.1134 0.0359</td>
<td>0.0973 0.0395</td>
<td>0.0600 0.0460</td>
<td>0.0797 0.0351 0.0645 0.0424</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-28917</td>
<td>-28879</td>
<td>-28855</td>
<td>-18643</td>
</tr>
<tr>
<td>N x T</td>
<td>15385</td>
<td>15385</td>
<td>15385</td>
<td>10136</td>
</tr>
</tbody>
</table>
Table 2.a: Baseline parameter estimates for property convictions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0109 0.0095</td>
<td>0.0225 0.0077</td>
<td>0.0062 0.0120</td>
<td>0.0292 0.0108</td>
<td>0.0469 0.0099</td>
<td>0.0386 0.0165</td>
</tr>
<tr>
<td>Unemployment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls for trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-</td>
<td>-</td>
<td>-0.5124 0.6024</td>
<td>-</td>
<td>-</td>
<td>-0.3090 0.8384</td>
</tr>
<tr>
<td>Inflation</td>
<td>-</td>
<td>-</td>
<td>0.0054 0.0163</td>
<td>-</td>
<td>-</td>
<td>0.0219 0.0200</td>
</tr>
<tr>
<td>Wage rate</td>
<td>-</td>
<td>-</td>
<td>1.0111 1.4780</td>
<td>-</td>
<td>-</td>
<td>0.6726 1.9507</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>-</td>
<td>-</td>
<td>0.4267 0.2028</td>
<td>-</td>
<td>-</td>
<td>0.1939 0.5619</td>
</tr>
<tr>
<td>Conviction rate</td>
<td>-</td>
<td>-</td>
<td>0.0752 0.3448</td>
<td>-</td>
<td>-</td>
<td>0.0055 0.4528</td>
</tr>
<tr>
<td>Prison population</td>
<td>-</td>
<td>-</td>
<td>0.2352 0.1655</td>
<td>-</td>
<td>-</td>
<td>0.1612 0.3380</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls for Family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 12-17</td>
<td>-</td>
<td>1.1939 0.2507</td>
<td>1.8228 0.3771</td>
<td>-</td>
<td>1.4650 0.3188</td>
<td>1.8134 0.4626</td>
</tr>
<tr>
<td>Age 18-24</td>
<td>-</td>
<td>1.8637 0.2620</td>
<td>2.2813 0.3648</td>
<td>-</td>
<td>1.7397 0.3335</td>
<td>2.1322 0.4451</td>
</tr>
<tr>
<td>Age 25-34</td>
<td>-</td>
<td>0.9666 0.2522</td>
<td>1.3173 0.3469</td>
<td>-</td>
<td>0.8014 0.3139</td>
<td>1.1978 0.4033</td>
</tr>
<tr>
<td>Age 35-44</td>
<td>-</td>
<td>0.3556 0.2689</td>
<td>0.3845 0.3024</td>
<td>-</td>
<td>0.5855 0.3129</td>
<td>0.6380 0.3422</td>
</tr>
<tr>
<td>Ratio males</td>
<td>-</td>
<td>1.3316 0.3848</td>
<td>1.4821 0.3898</td>
<td>-</td>
<td>1.2240 0.3950</td>
<td>1.2460 0.6141</td>
</tr>
<tr>
<td>SES</td>
<td>-</td>
<td>-0.0681 0.0171</td>
<td>-0.0150 0.0174</td>
<td>-</td>
<td>-0.3128 0.0222</td>
<td>-0.2848 0.0268</td>
</tr>
<tr>
<td>Statistical controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-5.1474 0.0982</td>
<td>-6.5314 0.5615</td>
<td>-7.1055 0.6116</td>
<td>-5.2898 0.1044</td>
<td>-5.8947 0.7243</td>
<td>-6.2504 0.7707</td>
</tr>
<tr>
<td>State dependence</td>
<td>0.1902 0.0205</td>
<td>0.2101 0.0201</td>
<td>0.1957 0.0207</td>
<td>0.1642 0.0216</td>
<td>0.1692 0.0213</td>
<td>0.1667 0.0214</td>
</tr>
<tr>
<td>Family-specific</td>
<td>1.0067 0.0824</td>
<td>0.9564 0.0795</td>
<td>0.9595 0.0799</td>
<td>0.9795 0.0846</td>
<td>0.9492 0.0825</td>
<td>0.9590 0.0824</td>
</tr>
<tr>
<td>Time-varying</td>
<td>0.1645 0.0428</td>
<td>0.0241 0.0654</td>
<td>0.0000 0.0475</td>
<td>0.1090 0.0393</td>
<td>0.0204 0.0459</td>
<td>0.0094 0.0388</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-28082</td>
<td>-28037</td>
<td>-28024</td>
<td>-17946</td>
<td>-17917</td>
<td>-17912</td>
</tr>
<tr>
<td>N x T</td>
<td>15385</td>
<td>15385</td>
<td>15385</td>
<td>10136</td>
<td>10136</td>
<td>10136</td>
</tr>
</tbody>
</table>
## Table 3.a: Baseline parameter estimates for violent convictions

<table>
<thead>
<tr>
<th></th>
<th>1920-2005</th>
<th>1950-2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violent Convictions</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unemployment</strong></td>
<td>0.0280</td>
<td>0.0178</td>
</tr>
<tr>
<td><strong>Controls for trend</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Inflation</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Wage rate</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Conviction rate</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Prison population</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Controls for Family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 12-17</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age 18-24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age 25-34</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age 35-44</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SES</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Statistical controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-6.2285</td>
<td>-6.7839</td>
</tr>
<tr>
<td>State dependence</td>
<td>0.3095</td>
<td>0.3046</td>
</tr>
<tr>
<td>Family-specific</td>
<td>0.9411</td>
<td>0.8961</td>
</tr>
<tr>
<td>Time-varying</td>
<td>0.1375</td>
<td>0.0388</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-25961</td>
<td>-25948</td>
</tr>
<tr>
<td>N x T</td>
<td>15385</td>
<td>15385</td>
</tr>
</tbody>
</table>
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

5.4 Time-varying unemployment effects

Next we examine whether the effect of the unemployment rate on the conviction outcomes changes significantly between 1920 and 2005. In Section 5.3 we found that the effect of the unemployment rate is higher for the period 1950-2005 compared to the period 1920-2005. More specific, all estimates for the models including all control variables for the 1920-2005 period were found insignificant. This was found for all three conviction samples. The largest increase between the two sample periods was found for property convictions. Despite these results, the choice for the sample split at year 1950 remains quite arbitrary. It was mainly determined by the visual analysis of Figure 5.2.

To formally investigate whether the effect of the unemployment rate is increasing, we estimate a time-varying effect for the unemployment rate. Time-varying regression models are discussed in detail by Harvey (1989). To implement the time-varying effect we decompose $\delta$ in equation (5.4) into a deterministic, family-specific and time-varying component. In particular

$$\delta = \delta^0 + \delta_i + \delta_{\tau t}, \quad \delta_i \sim N(0, \sigma^2_\delta), \quad \delta_{\tau t} = \delta_{\tau t-1} + \epsilon_{\tau t}, \quad \epsilon_{\tau t} \sim NID(0, \sigma^2_\epsilon), \tag{5.7}$$

where $\delta^0$ is the deterministic effect, $\delta_i$ is the family-specific and $\delta_{\tau t}$ is the time-varying effect. The deterministic and family-specific effect are modeled as in equation The time-varying effect $\delta_{\tau t}$ is modeled by a random walk process, with a normally distributed disturbance term $\epsilon_{\tau t}$. The disturbance term has mean zero and variance $\sigma$. To achieve identification we fix $\delta_{\tau 1} = 0$. Appendix C discusses the estimation procedure for the binomial panel data model with family-specific and time-varying random effects for the unemployment rate.

5.4.1 Results

The parameters of the binomial dynamic panel data model with a time-varying effect for the unemployment rate are estimated for the three conviction samples. We only estimated the model for the entire sample period 1920-2005. The estimates for the time-varying effect for the unemployment rate, $\delta_{\tau t}$ are visually displayed in Figure 5.4. For serious and violent convictions no significant changes in the effect of the unemployment rate are found. The estimated changes in the time-varying effect never
exceed above 0.2 percentage points. This confirms the findings in Section 5.3, where we found no significant overall effect for the unemployment rate for serious and violent convictions. Figure 5.4 shows that these findings are not resulting from an unlucky, or lucky, sample split.

The estimated time-varying effect for property convictions increases significantly between 1920 and 2005. Moreover the overall magnitude of the increase between 1920 and 2005 is between 2 and 6 percentage points. This indicates that during the sample period the families have become more responsive to changes in the unemployment rate with respect to property crimes.

Essentially this result formally confirms Figure 5.3, in establishing that the effect of the unemployment rate was insignificant prior to approximately 1960 and becomes significant after this period. The main question that remains is why this increase occurred. Or when formulated differently, why was there little response to the unemployment rate during the period of mass unemployment in the 1930’s.

Several possible explanations spring to mind, but are hard to test. Firstly, it can be argued that during the Great Depression, poverty was so widespread in the Netherlands, that very few objects to steal were around. Objects stolen before that era were much more often than now food items, such as eggs or windfall apples, that incurred hefty penalties. Such thefts by nature occurred and could take place more easily in more agrarian areas, and not in the inner city slums where the unemployed were more concentrated. Indeed, a number of US authors have shown that little evidence exists for a link between the widespread unemployment during the Great Depression and crime. The same was stated for England too, for an overview see (Huzel, 1986). For completeness reasons, it must be stated that several authors have found a positive association between unemployment and crime levels during this period too. All in all, the evidence is contradictory.

One explanation for the emergence of significance just before 1960 may be the increasingly widespread availability of luxury goods that were tradable, relatively light (and thus easy to carry) and not easily traceable to the legal owner: the transistor radio appeared in the Netherlands in 1957. Before that period, burglars had to transport items such as Persian carpets, or family silver. Several other explanations spring to mind, and even though they cannot be regarded as more than musings, and even though some are perhaps too abstract to be at all testable, they may provide a starting point.
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

to think about explanations.

Since the early sixties, citizens’ respect for the 'state' has decreased. In the sixties, with hippies, drugs, protests against the Vietnam war, against nuclear arms, the state became increasingly portrayed as the State and as the Enemy. This pervaded general thinking about prosecution and the criminal justice system as well. Criminal law theorizing was abolitionistic, the state and the criminal justice system should refrain from messing around in people’s private lives. Prosecution was extremely restrained (‘dismiss, unless’) and sentences were low. Thus, one could argue, people knew they would likely not be prosecuted and if so not be punished severely, and the 'blemish' of arrest might even flip to become an heroic act (‘proletarisch winkelen’ [proletarian shopping = thieving]) in some circles (‘die pet pakt ons allemaal’ ['the copper’s hat nicks us all’ as a rewording of ‘that copper’s hat fits us all’).

Another explanation might be the advent of drugs and hard drugs around that time, and the quick appearance on the scene of junkies, who had to steal extensively if they had no source of income to finance their drug habit. In the 1970’s the Netherlands has the lowest incarceration rate in Europe (and the lowest ever recorded for the Netherlands), and possibly it is simply the case that very few criminals are incapacitated, and thus are able to steal.

A methodological problem for our series is that there is some volatility in definitions of unemployment in the 1980s as successive governments tried to show the success of their policies by definitional fiddling.

Crime and the perception and role of the criminal justice system changed over the years too. As crime levels increased, the public and authorities became increasingly fed up with insecurity and public disorder, and the discourse shifted to a victim-oriented one: the state should protect victims, not norms or the social order. By the early 90s, the Dutch state starts building prisons to make up for lacking prison space (in the 80s a number of scandals showed that prison space was so scarce that sex offenders could not be housed and were sent home through the back door). Simultaneously, possibly to cope with rising numbers of arrestees and defendants, prison sentences became increasingly reserved for violent offenders and drugs offenders. The ‘officiersmodel’ whereby also a prosecutor could impose non-custodial sentences, meant that you could essentially plea bargain your theft case so that you would not have a criminal record and came away with a 'taakstraf' (community sanction or a fine. Soon after the so-called 'Wet
Mulder' came into force where drunk driving and speeding (up to a certain severity) became something you 'bought' - you simply got an 'acceptgiro' (bank transfer card) at home and you paid up. Thus, a prosecution for a non-serious crime (such as theft or speeding) became not a transgression against the social order that carried the risk of incarceration (in the 1970s many prison sentences were for drunk driving) but a nuisance that was simply costly - and that you could of course steal for again. This fed into the changing perception of the state and politicians who were less and less seen as respectable old men who knew what is good for the citizen, but as a set of crooks (as opposed to 'politically wrong' in the sixties): the public discourse and discontent reflected thoughts along the line of: 'well the government robs me and cheats me so I can rob them/ cheat them'.

By the early 1990s, the welfare state started to break down. Unemployment became increasingly costly for the state and it became less easy for citizens to get welfare support. The allowances were tightened, enforcement of rules was increased, fraud was actively targeted. Especially for juveniles it became less easy to get social benefits. The allowance of those who did not apply for jobs frequently enough could be reduced, or withheld altogether. The government now requested some effort from the unemployed - contrary to the previous period - and welfare support was neither sure nor generous. Also, governmental policies to aid the unemployed were now restricted to the long-term unemployed only: the recently unemployed had to fend for themselves. All in all, one could argue that the unemployed were having a tougher time making ends meet from this period onwards.

5.5 Conclusion

In this paper we estimated the effect of the national unemployment rate on the serious, property and violent convictions time series of 181 families from The Netherlands. Two analysis were performed and discussed. First, we showed for all families that only property conviction time series starting after 1950 responds to changes in the unemployment rate. This result was consistent for models that differed with respect to the inclusion of control variables. Second, a time-varying effect for the unemployment rate was estimated. The results clearly showed that the effect of the unemployment rate on property crimes is increasing between 1920 and 2005. The changes become
CHAPTER 5: THE EFFECT OF UNEMPLOYMENT ON CRIME

Figure 5.4: The estimated time-varying effects for the unemployment rate.

significant after approximately 1960.
In this appendix we discuss the evaluation of the likelihood for the binomial panel data model, given by equations (5.1), (5.4), (5.5) and (5.6). An exact exposition is presented in Mesters and Koopman (2014). The likelihood is defined as 
\[ \ell(\psi) = \log p(y), \]
where \( y = \{y_{i,\tau}\}_{i=1,...,N, \tau=1,...,T}. \) From the nonlinearity of the observational density and the presence of the random family-specific and time-varying components it follow that the likelihood does not exists in closed from. Nonetheless, the joint density of the observations can be expressed as
\[ p(y) = \int_{\theta} p(y, \theta; \psi) d\theta = \int_{\mu} \int_{\xi} p(y, \mu, \xi; \psi, x, w) d\mu d\xi, \] \hfill (5.8)

where \( \mu = (\mu_1, \ldots, \mu_N)' \) is the vector of family specific effects, \( \xi = (\xi_1, \ldots, \xi_\tau)' \) is the vector of time-varying effects, \( x = \{x_{i,\tau}\}_{i=1,...,N, \tau=1,...,T} \) is the collection of family control variables and \( w = (w_1, \ldots, w_\tau)' \) is the vector of common control variables. The joint density can be rewritten as
\[ p(y) = \int_{\mu} \int_{\xi} p(y|\mu, \xi; \psi, x, w) p(\mu)p(\xi)d\mu d\xi, \] \hfill (5.9)

where \( p(y|\mu, \xi; \psi, x, w) \equiv p(y|\theta; \psi) \), which is given in equation (5.3) which follows from the independence between the family-specific and the time-varying random effects. To evaluate the high-dimensional integral in (5.9) efficiently we use the importance sampling technique. An importance sampling representation for (5.9) is given by
\[ p(y) = g(y; \hat{\xi}) g(\mu; \hat{\mu}) \int_{\xi} \int_{\mu} \frac{p(y|\mu, \xi; \psi, x, w)}{g(y|\mu; \xi) g(\xi|\psi; \hat{\mu})} g(\mu|y; \hat{\xi}) g(\xi|y; \hat{\mu}) \, d\mu \, d\xi, \] \hfill (5.10)

where \( g(\mu|y; \hat{\xi}) \) and \( g(\xi|y; \hat{\mu}) \) are the importance densities. We define \( \hat{\mu} \) and \( \hat{\xi} \) as the posterior modal values of \( p(\mu, \xi|y; x) \), that is \( \{\hat{\mu}, \hat{\xi}\} \equiv \arg\max_{\mu,\xi} p(\mu, \xi|y; x) \). The choice for the posterior modal values is not necessary as any sufficient statistic can be used for the conditioning. It is used for computational convenience.
modal values and the importance densities the joint density can be estimated by

\[
\hat{p}(y) = g(y; \hat{\xi})g(y; \hat{\mu}) \sum_{i=1}^{M} \frac{p(y|\mu^{(i)}, \xi^{(i)}; \psi, x, w)g(y|\mu^{(i)}; \xi)g(y|\xi^{(i)}; \hat{\mu})}{g(y|\mu^{(i)}; \xi)g(y|\xi^{(i)}; \hat{\mu})},
\]

(5.11)

where samples \(\mu^{(i)}\) and \(\xi^{(i)}\) are drawn from the importance densities \(g(\mu|y; \hat{\xi})\) and \(g(\xi|y; \hat{\mu})\), respectively.

Next we discuss the construction of the importance densities. We choose both densities to follow Gaussian distributions and modify their means and variances such that their modes are equal to the modes of the original posterior density \(p(\mu, \xi|y; \psi, x, w)\). So (2003) and Jungbacker and Koopman (2007) argue that this strategy can be implemented by numerically maximizing \(\log p(\mu, \xi|y; \psi, x, w) = \log p(y|\mu, \xi; \psi, x, w) + \log p(\mu, \xi) - \log p(y; \psi, x, w)\) with respect to \(\mu\) and \(\xi\). The instrumental basis to facilitate this numerical maximization is given, for variable \(y_{i,t}\), by the linear Gaussian panel data model

\[
y_{i,t} = c_{i,t} + \theta_{i,t} + u_{i,t}, \quad u_{i,t} \sim NID(0, d_{i,t}^{2}),
\]

(5.12)

where \(c_{i,t}\) is a fixed constant, stochastic component \(\theta_{i,t}\) is given by equation (5.4) and \(u_{i,t}\) is a random variable with mean zero and fixed variance \(d_{i,t}^{2}\). The constants \(c_{i,t}\) and \(d_{i,t}\) are chosen such that (5.12) can be used to compute the posterior modal values \(\hat{\mu}\) and \(\hat{\xi}\), respectively. The elements \(u_{i,t}\) and \(\theta_{j,s}\) are uncorrelated with each other, for all \(i, j = 1, \ldots, N\) and \(s, t = 1, \ldots, T\). Furthermore, \(u_{i,t}\) is serially uncorrelated. It follows that

\[
g(y|\mu, \xi) = \prod_{i=1}^{N} \prod_{t=1}^{T} g(y_{i,t}|\mu_{i}, \xi_{t}), \quad \text{with} \quad g(y_{i,t}|\mu_{i}, \xi_{t}) \equiv NID(c_{i,t} + \theta_{i,t}, d_{i,t}^{2}).
\]

(5.13)

The maximization of \(\log p(\mu, \xi|y; x)\) with respect to \(\mu\) and \(\xi\) can be carried out via the Newton-Raphson method. This procedure is summarized in the following algorithm.

**Algorithm A**

1. Initialize the algorithm by choosing \(\mu^{*}\) and \(\xi^{*}\) as starting values, which gives \(\theta_{i,t}^{*}\), for all \(i = 1, \ldots, N\) and \(t = 1, \ldots, T\);
2. Given the set of two equations

\[
\frac{\partial \log p(y_{i,t}|\theta_{i,t}; \psi)}{\partial \theta_{i,t}} = \frac{\partial \log g(y_{i,t}|\theta_{i,t})}{\partial \xi_{i,t}}, \quad \frac{\partial^2 \log p(y_{i,t}|\theta_{i,t}; \psi)}{\partial \theta_{i,t} \partial \theta_{i,t}} = \frac{\partial^2 \log g(y_{i,t}|\theta_{i,t})}{\partial \theta_{i,t} \partial \theta_{i,t}},
\]

for \( i = 1, \ldots, N \) and \( t = 1, \ldots, T \), where \( p(y_{i,t}|\theta_{i,t}) \) is the observation model (5.1) and \( g(y_{i,t}|\theta_{i,t}) \) is given by (5.13), we can deduce expressions for \( c_{i,t} \) and \( d_{i,t} \) as functions of \( \theta_{i,t} \), and compute \( c_{i,t} = c^*_{i,t} \) and \( d_{i,t} = d^*_{i,t} \) for \( \theta_{i,t} = \theta^*_{i,t} \);

3. Compute \( \tilde{\mu} = E_g(\mu|y; \xi^*) \) from the resulting model (5.12) with \( \xi = \xi^* \), \( c_{i,t} = c^*_{i,t} \) and \( d_{i,t} = d^*_{i,t} \);

4. Replace \( \mu^* \) by \( \mu^* = \tilde{\mu} \);

5. Compute \( \tilde{\xi} = E_g(\xi|y; \mu^*) \) from the resulting model (5.12) with \( \mu = \mu^* \), \( c_{i,t} = c^*_{i,t} \) and \( d_{i,t} = d^*_{i,t} \);

6. Replace \( \xi^* \) by \( \xi^* = \tilde{\xi} \);

7. Iterate from (ii) to (vi) until convergence.

Since the mode and the mean of the approximating linear Gaussian model are set equal to the mode of the original model, it holds that \( \tilde{\mu} = \hat{\mu} = \arg\max_{\mu} p(\mu|y; \hat{\xi}; x) \) and \( \tilde{\xi} = \hat{\xi} = \arg\max_{\xi} p(\xi|y; \hat{\mu}; x) \). Further, as \( \mu \) and \( \xi \) are independent, it holds that \( \{\tilde{\mu}, \tilde{\xi}\} = \arg\max_{\mu, \xi} p(\mu, \xi|y; x) \).

The computation of \( E_g(\mu|y; \xi^*) \) and \( E_g(\xi|y; \mu^*) \), in steps (iii) and (v), respectively, is handled by applying the Kalman filter and smoothing recursions to the approximating model. This procedure can be accelerated by first collapsing the approximating model twice. Once over the cross-section and once over the time series dimension, see Mesters and Koopman (2014) for further details. After the posterior model values are obtained and the approximating model is fitted, sampling \( \mu^{(i)} \) and \( \xi^{(i)} \) can be drawn from the importance densities using the simulation smoother methods of Durbin and Koopman (2002). The likelihoods of the approximating model \( g(y; \hat{\xi}) \) and \( g(y; \hat{\mu}) \) can be evaluated by the prediction error decomposition provided by the Kalman filter.
Appendix B

This appendix discusses the estimation of the binomial dynamic panel data model with heterogeneous slope effects, given by equations (5.1), (5.4), (5.5) and (5.6). For this model signal (5.4) can be rewritten as

\[ \theta_{i, \tau_t} = \delta^0 \text{UN}_{\tau_t} + x_{i, \tau_t} \beta + w_{\tau_t} \lambda + \gamma y_{i, \tau_{t-1}} + a'_{\tau_t} m_i + \xi_{\tau_t}, \]  
(5.14)

where \( a_{\tau_t} = (1, \text{UN}_{\tau_t})' \) and \( m_i = (\mu_i, \delta_i)' \). The family-specific effect \( m_i \) is here a vector consisting of the mean family-specific effect \( \mu_i \) and the family-specific effect of the unemployment rate \( \delta_i \). These effects are loaded to the signal by the partially time-varying loading vector \( a_{\tau_t} \). When replacing \( \mu_i \) by \( m_i \) in Appendix A the estimation procedure is carried out in exactly the same manner.

Appendix C

This appendix discusses the estimation of the binomial dynamic panel data model with heterogeneous and time-varying slope effects, given by equations (5.1), (5.4), (5.5), (5.6) and (5.7). For this model signal (5.4) can be rewritten as

\[ \theta_{i, \tau_t} = \delta^0 \text{UN}_{\tau_t} + x_{i, \tau_t} \beta + w_{\tau_t} \lambda + \gamma y_{i, \tau_{t-1}} + a'_{\tau_t} m_i + b'_{\tau_t} f_{\tau_t}, \]  
(5.15)

where \( b_{\tau_t} = (1, \text{UN}_{\tau_t})' \) and \( f_{\tau_t} = (\xi_{\tau_t}, \delta_{\tau_t})' \). The time-varying effect \( f_{\tau_t} \) is here a vector consisting of the mean time-varying effect \( \xi_{\tau_t} \) and the time-varying effect of the unemployment rate \( \delta_{\tau_t} \). These effects are loaded to the signal by the partially time-varying loading vector \( b_{\tau_t} \). Note that \( b_{\tau_t} \) is fixed and known for each year \( \tau_t \). When replacing \( \xi_{\tau_t} \) by \( f_{\tau_t} \) in Appendix A the estimation procedure is carried out in exactly the same manner.
Chapter 6

General conclusion

In this thesis we empirically investigated to what extent the characteristics of a disadvantaged childhood, which resulted in a period in a juvenile treatment facility, had lasting impacts on socioeconomic adult outcomes, and whether life course transitions, such as those from employment and intimate relationships, additionally altered adult life outcomes. We proposed a conceptual framework which hypothesized a developmental process to explain adulthood offending and a variety of other socioeconomic outcomes. The main components in the conceptual framework were childhood skills, childhood signals and dynamic structural relationships among adult outcomes.

The conceptual framework was translated step by step to an empirical framework which was subsequently tested using observational data for samples of disadvantaged youths. These youths were institutionalized in a juvenile treatment facility in The Netherlands in the early 1990s and cannot be regarded as representative for the general population in the Netherlands. The samples rather characterized the lives of individuals who live in close proximity to offending. Additionally, these disadvantaged youths have low employment participation rates and are often the recipients of social welfare payments.

In this conclusion we discuss our main empirical findings. In particular, we briefly summarize the most important findings from the different chapters and discuss the similarities and differences that were found across the chapters. Further, we discuss the implications of these findings for criminological life course theories. Based on these implications we outline some extensions for the theoretical perspectives which empha-
size a more dynamic role for skills. We note that the extensions are by no means intended as a complete theoretical perspective. Next to the theoretical implications we discuss the consequences from our findings for public policy. We differentiate between individual-level intervention policies and the more general role of government policies in explaining offending. The methodology that was adopted in this thesis is briefly compared to existing methods used in life course criminology. We finish with a discussion regarding possible directions of future research.

6.1 Main empirical findings

In the second chapter we documented to what extent childhood and adolescent skills determined adulthood offending outcomes for disadvantaged youths. The childhood skills included both cognitive and social skills and the adult offending outcomes were split into categories that contained serious, property and violent offenses. The chapter was framed to investigate to what extent dynamic mean-level effects from childhood and adolescent skills could explain the age-crime curve for serious, property and violent offending.

We found that childhood and adolescent skills have lasting effects on adult offending. This implies that individual differences in adolescent and adult offending can to some extent be explained by differences in skills that are obtained during childhood and adolescence. More important was the finding that the effects of the skills were not stable over the life course. In particular, effects from both cognitive and social skills changed in magnitude during adolescence and adulthood. This implies that different skills are important in explaining adult offending at different stages of life.

Cognitive skills, as measured by intelligence, were found important predictors for male offending after the adolescent period. During adolescence cognitive skills were less important, but differences in social skills were able to explain the peak in adolescent offending. For females social skills were found overall more important. The significant personality traits, that were used to proxy the social skills, were neuroticism, extroversion and thrill seeking. These traits are mainly important during the adolescence for both males and females.

In the third chapter we documented dynamic interactions between adulthood outcomes for offending, employment and social welfare. We treated the childhood as
CHAPTER 6: GENERAL CONCLUSION

an unobserved period, for which we statistically controlled. The chapter was framed to distinguish between economic and sociological perspectives for the relationship between employment and crime. We used social welfare as an identifying mechanism by arguing that social welfare only provides financial gains that are deemed important in economic theories, whereas social welfare does not provide the social bonds that are deemed important in sociological theories.

We found that a substantial part of the relationship between employment and crime was spurious. In particular, the statistical control variables for employment and offending were negatively correlated. This indicates that individuals who on average have a higher propensity towards criminal behavior also have a lower propensity towards employment. No attempts were made in this chapter to explain this unobserved part of the relationship, but as shown in chapter four at least part of the latent correlation could be explained by differences in skills.

Instead of explaining the unobserved part of the relationship between crime and employment, the focus was on the structural relationships between offending, employment and social welfare. We found significant bi-directional negative structural effects between employment and offending. This implies that employment reduces future offending and that offending reduces future employment. Some refinements of the model showed that only regular employment was able to reduce property offending, while both violent and property offending were shown to have large negative consequences for future employment probabilities.

The relationship between social welfare and offending was more subtle. Only the public assistance category of social welfare was able to significantly reduce property offending. Public assistance is the form of welfare that is based solely on financial need. The effect of this form of welfare on property offending was found similar in magnitude as when compared to the effect from employment on property offending. This highlights the important role for financial gains, and economic or rational choice theories, in explaining the structural part of the relationship between employment and property crime for disadvantaged youths.

In chapter four the insights from chapters two and three were combined. In particular, chapter two stressed the importance of childhood and adolescent skills in explaining adulthood offending and chapter three highlighted the interaction between adult outcomes for offending, employment and social welfare. The model in chapter four
combines the explanatory power of the childhood and adolescent skills while allowing for dynamic interaction between adult outcomes. The chapter was framed to determine during which stage (childhood or adulthood) investments might best be made to improve the adult life outcomes of disadvantaged youths. We studied a broader array of adult outcomes that included offending, employment, social welfare, drug use and intimate relationships. We additionally included childhood signals in our model to study the effects from education levels and criminal records on adult outcomes over and above the effects of the childhood skills.

The results suggest that cognitive and social skills have lasting effects on male adult outcomes for offending and employment. For females only social skills have lasting effects on offending and employment. For offending similar results were found compared to chapter two, even though in chapter four the skills were included using a factor model that incorporated more measurements for the skills. The childhood education signals are important for reducing social welfare during adolescence and increasing employment during adulthood. The criminal records signal increased adolescent and adult offending and drug use.

Next to the persistent effects from the childhood skills and signals we found many structural dynamic interactions between the adult outcome variables. The main implications concerning the structural relationships between offending, employment and social welfare that were developed in chapter three continued to hold in chapter four. Additionally offending was reduced by intimate relationships, whereas drug use increases the offending probability. For females intimate relationships were found to have more explanatory power when compared to males. In particular, intimate relationships predict reductions in social welfare during adolescence and they are negatively interlinked with drug use.

Simulated increases in cognitive and social skills showed that large reductions in the offending probabilities can be achieved. These reductions are persistent in the sense that they lower offending probabilities over the entire adolescent and adult life course. Overall social skills were found more important in reducing offending behavior. Additionally, and equally important, by increasing social skills employment probabilities also increased significantly during adulthood. For females, increases in social skills also significantly lowered the probabilities for intimate relationships and drug use during adolescence. Thus, increases in social skills lead to large gains on multiple adolescent
and adult outcomes for both males and females.

The findings from the fifth chapter can be regarded as a warning for developmental life course studies in general. While the chapter documented interesting empirical findings, from the perspective of the thesis it mainly stressed the dependence of interactions between adulthood outcomes on historical contexts. In particular, the chapter showed that for historically disadvantaged families the relationship between unemployment and property crime changed during the last century. Between 1930 and 1960 no significant relationship between unemployment and property crime was found, while from 1960 until 2005 an increasingly strong positive effect from unemployment on property crime was found. This finding is remarkable when realizing that during the great depression in the 1930s unemployment rates skyrocketed in the Netherlands. This finding suggests that caution is warranted when generalizing results from older samples to current situations. Similar changes in the relationship between crime and intimate relationships are found in Beijers, Bijleveld, and van Poppel (2012).

6.2 Implications for life-course theories

Several results from this thesis have implications for criminological life course theories. We briefly summarize which findings exactly agree or disagree with the dominant perspectives in criminology. To keep the discussion within reasonable length we only include reflections on the self-control theory of Gottfredson and Hirschi (1990), the age-graded theory of social control (Sampson & Laub, 1993; Laub & Sampson, 2003; Sampson & Laub, 2005) and the economic rational choice perspective (Becker, 1968; Ehrlich, 1973; Block & Heineke, 1975). Based on these reflections we suggest a number of extensions that mainly concern the role of skills in explaining offending.

In a seminal contribution Hirschi and Gottfredson (1983) state three provocative hypotheses concerning the relationship between age and crime. First, their invariance hypothesis states that the relationship between age and crime is largely invariant across cultures, demographics and other social contexts. Second, their non-interactive hypothesis states that the same variables are responsible for explaining the cross-sectional variation in offending at any age. The variables that they deem relevant are related to a higher order construct that they label self-control in their general theory of crime (Gottfredson & Hirschi, 1990). By result their general theory is often referred to as
self-control theory. Third, their inexplicability hypothesis argues that no combination of sociological and psychological variables can be constructed to explain the variation in offending that occurs with age. Only age itself can explain the variation.

The empirical results from this thesis conflict with the non-interactive hypothesis and the inexplicability hypothesis. We did not investigate the invariance hypothesis since our empirical results were generated only for specific subgroups of disadvantaged youths. We refer to Steffensmeier et al. (1989) for some reflection on this hypothesis.

Throughout this thesis we have consistently found that childhood skills, of which some may reflect aspects of self-control, have persistent effects on adulthood offending. In this sense, self-control appears to contribute to explaining offending over the life span. Despite this, our empirical results also showed that different variables were important in explaining offending in different periods. Both the effects from childhood skills and signals, and the interactions between adult outcome variables changed in magnitude and sometimes in sign over the life course. Not all the changes in effect sizes were significant, but a considerable number did suggest that different variables were important in explaining offending at different stages of life. Examples included the effects of social skills on offending and the effects of employment on offending. These findings provide evidence against the non-interactive hypothesis which suggests that the same variables consistently explain between-individual differences in offending over the life course.

Second, the inexplicability hypothesis is also largely rejected by our findings. We found that when allowing the payoffs from childhood skills, both cognitive and social skills, to vary with age, we could explain the aggregate variation in the empirical age-crime curves for disadvantaged youths. Moreover, other adult life outcomes, such as employment, social welfare and drug use, were able to significantly predict offending. Similar as in Sweeten et al. (2013) we must conclude that combinations of sociological and psychological variables are able to explain the variation in offending that occurs with age.

Overall, our empirical findings have little in common with self-control theory. Two comments are in order. First, our conceptual framework allowed for the possibility that childhood skills could explain the between-individual variation at all ages, but it was by no means intended as a direct test for self-control theory. Second, we do not claim to have adequately operationalized “self-control” in the sense of Gottfredson
CHAPTER 6: GENERAL CONCLUSION

and Hirschi (1990). We believe that certain elements of self-control are included in the skills, but that our separation of skills into cognitive and social skills is more useful for practical purposes.

The age-graded theory of social control posits a more dynamic explanation for offending over the life course (Sampson & Laub, 1993). It argues that different social bonds are important at different stages in life in explaining offending. In particular, during childhood attachments to parents are important, while during adolescence school and peers form important social bonds and finally during adulthood intimate relationships and employment may provide bonds that provide barriers to criminal behavior. Further, Sampson and Laub (1993) suggest the mechanism of cumulative disadvantage to explain persistence in offending behavior as offending can reduce the possibility of obtaining future social bonds leading to a negative spiral of disadvantage.

In Laub and Sampson (2003) their view for the role of childhood skills is outlined. In particular, they argue that aspects such as low IQ and aggressiveness can predict age-invariant between-individual differences with reasonable accuracy. But they reject the view that skills can explain different developmental trajectories in the sense of Moffitt (1993). Also, they do not investigate the possibility that skills can change over the life course, or that the effect of skills on offending can change over the life course.

The majority of the youths that we studied in this thesis were missing important social bonds during childhood and adolescence. They often came from disrupted family environments and during adolescence they were, at least temporarily, institutionalized in a treatment facility. Thus, their on average poor socioeconomic adult outcomes should not come as a surprise when viewed from the perspective of cumulative disadvantage. This perspective is further strengthened by the consistent finding that offending lowers future employment probabilities (chapters three and four).

Nevertheless, as illustrated throughout the thesis, considerable heterogeneity exists within the on average poor adult outcomes. When combining this finding with the significant dynamic effects from employment, social welfare and drug use on adult offending one might be inclined to regard this as evidence in favor of social control theory. Additionally, we found that several of these effects significantly change over the life course stressing the importance of the age-graded view. Despite this bundle of apparent evidence in favor of age-varying social bonds, all these findings can also be explained from an economic rational choice perspective (e.g. Ehrlich, 1973; Grogger,
In 1998). The argument goes as follows. Employment and social welfare provide financial gains that reduce the relative returns from criminal behavior. With maturation employment and social welfare may become more negatively associated with offending as a result of increasing financial demands. This reasoning provides an alternative explanation for the same empirical findings.

The rational choice perspective is somewhat strengthened, albeit not conclusively proven, by two other findings. First, the public assistance category of social welfare was able to significantly reduce property crime with a coefficient that was nearly the same in magnitude as the effect that was found for employment on crime. This indicates that at least for property related offenses the financial aspect from employment as opposed to the social aspect is important. Second, we found that intimate relationships, which do not provide financial gains, do not significantly reduce offending.

Overall, the perspective of cumulative disadvantage is confirmed by the empirical results, but the social bonds aspect remains hard to distinguish from the economic rational choice perspective. Interestingly, both self-control, social control and economic rational choice theories assign limited importance to skills.

6.2.1 An age-graded perspective for skills

Next, we argue for a more dynamic perspective for skills in explaining adult offending. Our suggestions are based on the empirical findings from this thesis, the economic model for childhood skill formation of Cunha and Heckman (2007) and the developmental perspective on personality development (Caspi et al., 2005). We follow Sampson and Laub (2005) and retain the notion that persistent offending and desistance, and hence trajectories of crime, can be meaningfully understood within the same theoretical framework.

Our suggestions concern two aspects. First, we view abilities, personality traits, motivations and skills as dynamic concepts that can be altered by investments over the life course. For the present exposition we label these factors as skills, while acknowledging that this broad class contains different components. Skills can be increased or decreased up to a set of stabilizing ages that depend on the individual skills. Second, given the dynamic conceptualization of skills and the developmental stage that is theorized (childhood, adolescence, or adulthood) it is possible to determine which kind of
variation in skills needs to be addressed in life course criminological theories.

We discuss three levels of variation in skills that are potentially useful to distinguish between in order to explain offending over the life course. First, rank-stability suggests that while there is variation in skills over age, the ordering of the individuals within a certain population does not change with age. Throughout this thesis we have generally assumed rank-stability and evidence in Roberts and DelVecchio (2000) suggests that this is reasonable for many skills. This concurs with the finding that skills are, at least partially, able to explain cross-sectional differences in offending (Laub & Sampson, 2003).

Second, rank-order stability does not imply mean-level stability, where mean-level change refers to change in the quantity or level of some skill over time and is defined for a certain population. Roberts, Walton, and Viechtbauer (2006) review the literature on mean-level change in skills. When assuming that the majority of individuals in a population change in the same direction, mean-level change reflects normative alterations in the average amount of a skill in a population over time and may reflect maturational or historical processes common to that population (Blonigen, 2010). Closely related to mean-level change is the change in effect that a skill has on offending. More specifically, when skills stabilize with maturation this does not imply that the effect that skills have on offending also stabilizes. Interaction with age may remain to imply different effects from different skills over the life course. Distinguishing between mean-level changes in skills and age-varying effects of skills on offending requires longitudinal data for the skills. In this thesis we have investigated age-variation in effects of skills, which may to some extent may be confounded with mean-level changes skills.

Third, individual-level change suggests that skills can change for each individual with age. The most active period of individual-level change typically takes place during childhood (Cunha et al., 2006; Cunha & Heckman, 2007). The point where individual-level change stops depends on the particular skill that is considered. As mentioned before, cognitive abilities are typically found stable after age 10, whereas various social skills can vary on the individual-level until the early 20s (Hopkins & Brecht, 1975; Dahl, 2004). Also, there might be specific skills that can be developed well into old age. Thus, when studying the effect of skills on offending it is important to determine whether the skills can be regarded as stable.

Overall, life course criminological theories typically impose rank-order stability,
but do not account for mean-level changes and individual-level change. This thesis has empirically shown that mean-level change can be important in explaining aggregate offending rates within a population of disadvantaged youths. Cunha and Heckman (2008), Cunha et al. (2010) and many others provide evidence for individual-level change. Moreover, they provide evidence that investments in these skills by parents and others can significantly change the level of these skills.

6.3 Implications for public policy

In this section we discuss the implications for public policy that can be drawn from the main empirical findings. We distinguish between individual-level interventions and the role for more general government policies. The discussion concerning the individual-level interventions is based on the results from chapter four, whereas the implications for general government policies emphasize the role of social welfare.

6.3.1 Individual-level interventions

The main question for public policy that we aimed to answer in this thesis was at which stage during life (childhood or adulthood) investments, or interventions, can best be made to improve the adult life outcomes of disadvantaged youths. In principle a large number of interventions can be envisioned. We restrict ourselves to interventions that aim to improve childhood skills and interventions that facilitate employment. These individual-level interventions can naturally also improve society in general as they may lead to crime reductions and increase employment participation. Thus, the “costs” that are inevitably associated with such interventions need to be calculated within this broader perspective.

We discuss our main policy implications in the light of Figures 6.1 and 6.2. These figures summarize the results from chapter four, which include the insights from chapters two and three. We show the results from two hypothetical interventions for the male disadvantaged youths. For females similar results are shown in chapter four and the minor differences between males and females are not relevant for the implications. We emphasize that these results are merely thought experiments based on the estimated parameters.
CHAPTER 6: GENERAL CONCLUSION

The first intervention that is plotted suggests to improve social skills before age 16 by one standard deviation from their estimated mean. The implied consequences for offending from this intervention are shown by the squared dots in Figure 6.1 and for employment in Figure 6.2. The results show that this intervention significantly decreases offending over the entire adult life span. Also, employment is increased significantly from age 22 onwards. We emphasize that the estimates are based on the model from chapter four which includes childhood skills, childhood signals for education and criminal records, and adult life interactions between offending, employment, social welfare, drug use and intimate relationships. The main insight from this intervention is that it is a lasting intervention. It does not lead to rapid improvements in the lives of the disadvantaged youths, but generates large structural improvements that continue to last during adulthood on several domains.

The second intervention that we investigated was to create employment for individuals just after they turned 22. The results are shown by the triangular dots in Figures 6.1 and 6.2. We find that the predicted serious offending rates decrease rapidly and predicted employment rates increase rapidly. The gains are large in the first two years, but after this period they are not significant anymore. The reason for this is that these interventions do not solve the underlying problem, which is a lack of skills that enables these youths to retain a job for a prolonged period. We emphasize that the model does acknowledge that individuals learn from employment, however this does not outweigh the cumulative benefits from childhood skill improvements.

Several other arguments in favor of interventions that aim to improve social skills over interventions that facilitate employment can be found in the literature. First, Cunha and Heckman (2008) show that improvements in social skills can subsequently increase cognitive skills within the early childhood period. This can be regarded as an argument in favor of early childhood interventions as their results pertain to children below age 13 and cognitive skills are formed at early age (Hopkins & Brecht, 1975). Second, when arguing that birth is a random event in the sense that children have no influence on it, one can argue that the disadvantaged youths are to some extent not responsible for their disadvantaged childhood. Would it not be more “fair” to equalize the level of skills between children at young age instead of waiting until they are eligible for the job market. Third, the evidence for job market programs that facilitate employment is weak and inconclusive (Hill et al., 2011; Heckman & Kautz,
Fourth, there exists compelling evidence that investments in social skills are cost efficient (Heckman, 2006; Cunha & Heckman, 2007; Heckman & Kautz, 2013). This follows as these investments mutually reinforce themselves over the life span leading to accumulating benefits. This is also stressed by the role of adulthood multiplier effects in chapter four and is as such reflected in Figures 6.1 and 6.2.

The important question that remains is whether the hypothetical interventions that we showed are feasible. More specifically, is it possible to raise the level of skills in the way that we suggested. While this is not within the scope of the thesis we draw from the recent reviews of Hill et al. (2011) and Heckman and Kautz (2013) to point out a few interventions that have shown promising results. Overall, the interventions that they identify as promising are in line with the findings from the two hypothetical interventions shown in Figures 6.1 and 6.2. That is, social skill improvement programs are found effective and the evidence for job market programs that facilitate employment is weak and inconclusive.

First, recent research in economics has reanalyzed the results from infant and preschool programs, such as the Perry preschool program and the Abecedarian Program (Heckman, 2006; Heckman et al., 2010). These programs are typically intense programs that focus on multiple areas, such as morning school, parents and nutrition. While the long term results for improvements in cognitive ability were mixed, the improvements in social skills were large and conclusive. Further, similar as found in this thesis they found large improvements in adult outcomes on various domains that can be attributed to improvements in social skills. Second, after-school programs, social skills training and family interventions have all shown promising results for improving social skills (Hill et al., 2011). However, the largest improvements are found when the interventions occur on multiple areas. Third and finally, there are also interventions that were identified as ineffective. Examples include, juvenile awareness programs, boot camps and incarceration. The latter finding is also strongly reflected in the empirical results of this thesis. The juvenile criminal record that the disadvantaged youths carry into adulthood has predictive power for adult offending and drug use. Also, incarceration was shown to lower future employment probabilities (chapter three).

The interventions that were reviewed by Hill et al. (2011) and Heckman and Kautz (2013) were mainly conducted in the United States. This raises the question whether similar interventions would also work in The Netherlands. In The Netherlands in-
CHAPTER 6: GENERAL CONCLUSION

Figure 6.1: Results for males for serious offending from hypothetical interventions.

Intervention programs may ask for recognition from the “Erkenningscommissie”. This entity then judges the intervention and its evaluation on a number of areas. A major issue with the methodology used for judging interventions is that it is not based on randomized experimental evidence. In fact, recognition is in practice mainly given based on theoretical considerations and quasi-experimental designs. This makes it difficult to determine which interventions should be adopted in The Netherlands. At the same time it suggests the need for more rigorous experimentation and evaluation for interventions in The Netherlands.

In conclusion, our hypothetical intervention that improved social skills during childhood leads to large structural gains on multiple adult outcomes, which exceed the gains from creating employment for individuals after adolescence. Certain real life interventions that aimed to improve social skills, which were evaluated by randomized designs in the United States, have shown promising results and make the case that this approach for improving the lives of disadvantaged youths is feasible.
6.3.2 More on government policies

Next to the individual-level interventions we also studied the role of social welfare in explaining adolescent and adult offending. The welfare system is in The Netherlands designed to reduce inequality by redistributing income in order to avoid poverty and social exclusion. Next to its redistributing function welfare policies may yield externalities on other adult life domains. Thus, when designing welfare policies it is important to learn as much as possible about the externalities, such that the “true” costs of social welfare can be accurately assessed.

The main findings in chapters three and four provide information for the role of social welfare in explaining offending behavior. Overall, from age 16 onwards social welfare significantly lowers the probability for serious offending for males and females. The majority of this effect is attributed to the public assistance category of social welfare and the property crime category of offending (chapter 3). Also, and perhaps
CHAPTER 6: GENERAL CONCLUSION

surprising, increases in social skills suggest that between ages 17 and 21 disadvantaged youths will participate more in social welfare programs (Figures 4.10 and 4.11). The latter finding may seem counter-intuitive, because when skills increase should these youths not depend less on welfare?

We argue that given their disadvantaged childhood this is entirely rational and should also be stimulated from a government perspective. In particular, since the majority of youths obtain only low degrees of education and typically, after leaving the juvenile treatment, start working immediately, they could instead use the financial support from social welfare to improve their education levels. This form of conditional welfare can lead to increases in education levels which subsequently increase employment participation and lower offending rates (chapter four). Thus, while these policies require investments with uncertain outcomes on average they are likely to earn themselves back.

Overall, we conclude that social welfare is an important government policy that does not only provide subsistence but also leads to reductions in offending.

6.4 Comparison to traditional modeling approaches

In this section we briefly compare the modeling approach that is adopted in this thesis to more conventional modeling approaches that are used in life course criminology. In particular, we reflect on the semi-parametric group-based trajectory modeling approach (Nagin & Land, 1993; Nagin, 2005), fixed effects and random effects panel data methods (Baltagi, 2005) and treatment-effect estimation methods.

A large number of studies in life course criminology decompose panels of individual-level offending outcomes into latent trajectories (Nagin, 2005). This approach assumes that the heterogeneity between individuals and with age can be approximated by a finite number of trajectories. More specifically, it assumes that there exists a discrete number of offender types. In contrast, in the predominant part of this thesis we approximate the unobserved heterogeneity in offending panels by continuous distributions. Several comments are in order. First, evidence from Keane and Wasi (2013) suggests that continuous distributions for capturing unobserved heterogeneity provide a better fit to the data. Also, there exists evidence that latent class approaches understate the heterogeneity that exists in panels (Elrod & Keane, 1995). Second, it is
questionable whether the offending trajectories from latent class methods are causally distinct (Sampson & Laub, 2005). The trajectories modeling approach implies that individuals are locked in a specific trajectory and changes to different trajectories, via for instance life course transitions, are generally not accommodated. Third, the rigorous polynomial form of the trajectories offers less flexibility when compared to the autoregressive processes adopted in this paper. Also, as noted in chapter two, ex-post analysis of the smoothed autoregressive processes can still be used to investigate commonalities in the model residuals.

Overall our approach is quite distinct from the latent class approach and it can best viewed as an extension of the random effects panel data modeling approach (Baltagi, 2005). The choice for random effects instead of fixed effects is a consequence of the non-linearity in the observed outcomes. While fixed-effects approaches require less stringent assumptions for the disturbances the incidental parameter problem cannot be solved for general nonlinear models (Neyman & Scott, 1948; Lancaster, 2000). When departing from the logistic random effects panel data model in Baltagi (2005, Chapter 11) our proposed empirical framework adds time-varying random effects (Durbin & Koopman, 2012) and factor structures that incorporate the skills. The combination of these extensions required us to adopt non standard estimation methods that were based on, and sometimes extended the methods developed in Mesters and Koopman (2014). These methods allow for the treatment of (a) more general forms of unobserved heterogeneity, (b) large panels in both cross-section and time series dimensions, (c) nonlinear observational densities and (d) missing values.

The empirical results were derived from observational data for generally small samples. We did not make use of experimental data and the samples are too small to rely on exogenous shocks to identify treatment effects. Instead we relied on a structural modeling approach where the assumptions on the data generation process were explicitly stated (Keane, 2010; Wolpin, 2013). Ideally, when experimental data is available the framework can be used to identify causal treatment effects within a structural framework (Heckman & Vytlacil, 2005).
6.5 Directions for future research

A few important questions come to mind for future research. First, all the results that were developed in this thesis all pertain to disadvantaged youths. The empirical framework however is not specific for disadvantaged youths. In principle it can be applied to any sample and a comparison with a general population sample seems relevant. This would highlight differences between disadvantaged youths and the general population.

Second, a more careful analysis of which skills are important seems relevant. In chapter two we selected a number of personality traits to proxy the cognitive and social skills based on prior research, while in chapter four we combined a large number of measurements using a factor model to proxy the cognitive and social skills. Both approaches have their own benefits and weaknesses, but it is clearly important to determine more specifically which skills are important for which outcomes. This information could then be used for designing programs that aim to improve skills. A more elaborate separation of “skills” could separate them into four categories that summarize cognitive abilities, personality traits, motivations and narratives (Roberts & Wood, 2006).

Third, throughout this thesis we have assumed that the effects from adulthood life transitions are independent from the childhood skills. It is questionable whether this is realistic. When arguing that individuals with more skills would benefit more from certain transition this would imply that individual-level heterogeneity exists in the structural effects. While the average effects presented in this thesis remain valid a distributional perspective for the effects of social skill improvements seems relevant. This could highlight particular individuals that would gain most from improvements in skills and specific life transitions.

Fourth, several extensions for the empirical framework can be envisioned. For example, health outcomes could be included which give another domain on which large benefits could be achieved (Campbell et al., 2014). Also, the childhood skill model of Cunha et al. (2010) could be appended to our model to gain more insight in the development of skills during childhood. Finally, the model can be considered for multiple generations and as such determine whether investments in childhood skills for disadvantaged youths have the potential to reduce inequality within societies.
References


the Market Structure in Panel Data. *Journal of Marketing Research, 32*, 1-16.

(GESIS Study-No. ZA4758, Version: 1.1.0, doi:10.4232/1.10155)


174


177
Welfare: an individual-level study on disadvantaged males.
(unpublished: see www.geertmesters.nl for paper)


Wijkman, M., van der Geest, V., & Bijleveld, C. (2006). Delinquentie van op civielrechtelijke en strafrechtelijke titel behandelde jongeren (delinquency after treat-
ment in a juvenile justice institution under civil or criminal law). *Tijdschrift voor Familie- en Jeugdrecht, 1*, 18–27.


Samenvatting (Summary in Dutch)

In dit proefschrift onderzoeken we de volwassen sociaal-economische uitkomsten van jongeren die een deel van hun jeugd in een justitiële jeugdinrichting hebben doorgebracht. Deze jongeren hebben vaak lage cognitieve en sociale vaardigheden, een laag opleidingsniveau en hadden in sommige gevallen contact met justitie. Naast deze factoren maakte ook een slechte thuissituatie dat zij gedurende hun tienerjaren een periode in een residentiële setting behandeld werden. Het doel van dit proefschrift is om een conceptueel en empirisch raamwerk te ontwikkelen om de effecten vast te stellen van een aantal verschillende facetten, voor jongeren met een dergelijke “achtergestelde jeugd”, op uitkomsten gedurende het volwassen leven. We onderzochten of de cognitieve en sociale vaardigheden, het opleidingsniveau en contact met justitie de volwassen sociaal-economische uitkomsten beïnvloedden. Tegelijkertijd erkennen we dat transities gedurende volwassenheid de daaropvolgende volwassen uitkomsten kunnen beïnvloeden. De focus van dit proefschrift ligt op het verklaren van de uitkomst “criminaliteit” voor adolescenten en volwassenen, maar verschillende andere sociaal-economische uitkomsten worden ook onderzocht. In het bijzonder onderzoeken we ook arbeidsmarktnsuitkomsten, uitkeringen, drugsgebruik en intieme relaties.

Het scheiden van effecten uit de kindertijd en het volwassen leven is van theoretisch belang, maar het ontwikkelen van overheidsbeleid om uitkomsten voor achtergestelde jongeren te verbeteren vereist extra inzicht. Hiervoor is het belangrijk om te bepalen op welk moment welke interventies, in de vorm van investeringen in een positief volwassen leven, het meest effectief en efficiënt zijn. Er zijn vele interventies die in potentie het leven van achtergestelde jongeren kunnen verbeteren en dit proefschrift richt zich op een selectie van facetten waarop zulke interventies zouden kunnen aangrijpen. We zijn vooral geïnteresseerd in interventies die de cognitieve en sociale vaardigheden van jongeren kunnen verbeteren, en interventies die banen creëren. Efficiency van interventies
wordt hier bedoeld in een breed perspectief waarbij niet alleen monetaire aspecten van belang zijn, maar ook de maatschappelijke effecten van uitkeringen en drugs gebruik en criminaliteit relevant worden geacht (Cohen, 1998).

De ontwikkeling van een raamwerk dat in staat is om de sociaal-economische uitkomsten voor volwassenen te verklaren in termen van vroeggemeten factoren en andere uitkomsten gedurende het volwassen leven is een uitdagende taak. In dit proefschrift proberen we deze puzzel stap voor stap te benaderen. Het raamwerk is gebaseerd op inzichten uit de criminologie, sociologie, economie en psychologie. De theorieën die uit deze disciplines naar voren komen vormen het conceptuele raamwerk dat wordt vertaald naar een empirisch model in wiskundige formulering. Deze vertaling maakt het mogelijk om het conceptuele raamwerk te testen met behulp van observationele gegevens en econometrische methoden.

In de eerste stap ontwikkelen we een model waarin vaardigheden die zijn opgedaan gedurende de jeugd verklarende factoren vormen voor criminaliteit gedurende het volwassen leven. De vaardigheden uit de jeugd omvatten cognitieve vaardigheden en een breed scala aan sociale vaardigheden, die gemeten worden aan de hand van persoonlijkheidskenmerken. Hoewel de rol van cognitieve vaardigheden van oudsher van groot belang is in de sociale wetenschappen, heeft de rol van de sociale vaardigheden minder aandacht gekregen in het verklaren van crimineel gedrag (Hill et al., 2011). Het model dat we ontwikkelen onderschrijft de intuitief aansprekende stellingen dat meerdere vaardigheden belangrijk zijn voor het verklaren van criminaliteit en dat verschillende vaardigheden belangrijk kunnen zijn in de verschillende stadia in het leven (Cunha & Heckman, 2007).

Ten tweede wordt een model ontwikkeld dat expliciteert dat verschillende uitkomsten gedurende volwassenheid elkaar dynamisch beïnvloeden. In het bijzonder onderzoeken we in hoeverre crimineel gedrag, werk en uitkeringen elkaar beïnvloeden gedurende het volwassen leven. Crimineel gedrag kan de kans op werk verminderen en werk kan op zijn beurt crimineel gedrag reduceren (Lageson & Uggen, 2013). Dergelijke bi-directionele relaties moeten worden ontward om de zogeheten relatieve opbrengsten? te kunnen beoordelen van investeringen in het volwassen leven. Meer specifiek: als de interventie arbeid faciliteert dan kan dit een reducerend effect hebben op criminaliteit en uitkeringen, maar daarna kan deze reductie in criminaliteit weer leiden tot betere arbeidsmarktmogelijkheden. De Nederlandse welvaartsstaat biedt mensen de mogelijkheid
om, onder voorwaarden, een uitkering te krijgen. Door het gelijktijdig onderzoeken van criminaliteit, werk en uitkeringen, kunnen we de samenhang tussen deze keuzes onderzoeken. Dit maakt het mogelijk om twee zaken te onderzoeken. Ten eerste kunnen we dieper ingaan op de motivatie voor criminaliteit. Dit kan als volgt gezien worden. Gegeven dat uitkeringen alleen financiële middelen bieden en niet leiden tot meer sociale banden met de maatschappij, kan gesteld worden dat als werk en uitkeringen hetzelfde reducerende effect hebben op criminaliteit, dat dan geld een belangrijke motivatie is voor criminaliteit. Ten tweede, als uitkeringen criminaliteit verlagen, dan is dit een extra factor waarmee rekening moet worden gehouden als de kosten van uitkeringen worden bepaald.

In de derde stap combineren we de inzichten van de eerste twee stappen en stellen we dat als we de effecten van een achtergestelde jeugd op volwassen uitkomsten willen kwantificeren, dat dan de transities die zich gedurende het volwassen leven voordoen zelf ook moeten worden meegenomen. Overgangen tijdens het volwassen leven kunnen in deze zin fungeren als “multipliers” die de effecten van investeringen in jongeren verhogen of verlagen. Stel dat vroegtijdige interventies gericht op het verbeteren van sociale vaardigheden in staat zijn crimineel gedrag te verminderen tijdens de adolescentie, dan kan deze vermindering van criminaliteit de kans op werk gedurende het volwassen leven vergroten bovenop het marginale effect van de investering in sociale vaardigheden op de kans op werk. De ontwikkeling van dit dynamische conceptuele raamwerk, de vertaling hiervan naar een empirisch model en het testen van het empirische model vormen de belangrijkste bijdragen van dit proefschrift voor het verklaren van criminaliteit en andere sociaal-economische uitkomsten voor achtergestelde jongeren.

De vierde studie die is opgenomen in dit proefschrift biedt een historisch perspectief op de relatie tussen werkgelegenheid en criminaliteit. De belangrijkste vraag die we bestuderen is in welke mate het effect van werkloosheid op de criminaliteit is veranderd in de loop van de vorige eeuw. Hoewel deze vraag op zichzelf interessant is, dient deze in het kader van het proefschrift gezien te worden als een waarschuwing. Namelijk dat de historische context van belang is als men onderzoek doet naar verklaringen van criminaliteit. Het is namelijk de vraag of dezelfde jeugdige vaardigheden en levenslooptransities die belangrijk waren in het verleden net zo belangrijk blijven in de toekomst. Dit impliceert dat een voortdurende evaluatie nodig van relevante factoren voor het bouwen van interventies voor achtergestelde jongeren.
In dit proefschrift testen we het empirisch model met behulp van data die betrekking heeft op twee populaties van achtergestelde jongeren die geïnstitutionaliseerd waren in een behandelinstelling in Nederland. We maken gebruik van twee verschillende populaties: een populatie van mannen en vrouwen, die waren opgenomen in de jaren ’90, en een populatie van mannen die waren opgenomen van 1911-1914, en hun nazaten. De populaties beslaan een segment van de Nederlandse samenleving die oververtegenwoordigd is in de criminaliteitsstatistieken (Boendermaker, 1999). We verwijzen naar de jongeren als achtergesteld, terwijl we erkennen dat andere karakteriseringen als kwetsbare jongeren of hoog risico-jongeren evenzeer toepasselijk zijn. In totaal worden er ongeveer 4.000 jongeren jaarlijks geïnstitutionaliseerd in een straf- of jeugdzorginstelling in Nederland (CBS, 2013). Op basis van hun vroege contacten met justitie en/of hun gedragsproblemen, kunnen deze jongeren worden beschouwd als behorende tot een kansarme subgroep van jongeren die een hoog risico hebben voor het plegen van criminaliteit, het vaak lastig hebben op de arbeidsmarkt en relatief vaak de ontvangers zijn van sociale uitkeringen (van der Geest, 2011; Mesters et al., 2014; Verbruggen, 2014). Tijdens hun verblijf in de instelling worden ze behandeld voor hun gedragsproblemen en krijgen ze laaggeschoold onderwijs aangeboden. In hun late tienerjaren verlaten deze jongeren vertrekken meestal de instelling en en begin hun “volwassen ” leven. Gezien hun moeilijke jeugd hebben ze vaak moeite met deze transitie naar volwassenheid (Osgood et al., 2005). Ons doel is om aan te geven op wat voor soort kenmerken van deze jongeren interventies het beste kunnen aangrijpen om de volwassen sociaal-economische uitkomsten van deze jongeren te verbeteren.

Resultaten

De resultaten van dit proefschrift kunnen als volgt worden samengevat. In het tweede hoofdstuk vonden we dat de vaardigheden die waren opgedaan tijdens de kindertijd en adolescentie blijvende effecten hebben op criminaliteit gedurende het volwassen leven. Dit houdt in dat individuele verschillen in crimineel gedrag gedurende de volwassenheid tot op zekere hoogte kunnen worden verklaard door verschillen in vaardigheden die zijn opgedaan tijdens de kindertijd en adolescentie. Belangrijker was de bevinding dat de effecten van deze vaardigheden niet stabiel waren over de levensloop. In
het bijzonder bleek dat de grootte van effecten - van zowel de cognitieve als de sociale vaardigheden - veranderden tijdens de adolescentie en volwassenheid. Dit houdt in dat verschillende vaardigheden belangrijk zijn om mee te nemen voor het verklaren van criminaliteit voor verschillende levensfasen.

Cognitieve vaardigheden, gemeten door intelligentie, bleken belangrijke voorspellers voor criminaliteit van mannen na de adolescentie. Tijdens de adolescentie waren de cognitieve vaardigheden minder belangrijk, maar in deze periode waren verschillen in sociale vaardigheden in staat om de piek in delinquent gedrag van adolescenten te verklaren. Voor vrouwen bleken sociale vaardigheden in het algemeen belangrijker dan cognitieve vaardigheden. De belangrijkste persoonlijkheidskenmerken die werden gebruikt werden om de sociale vaardigheden te meten, waren neuroticisme, extraversie en spanningsbehoeftes. Deze eigenschappen zijn vooral belangrijk voor het verklaren van criminaliteit tijdens de adolescentie voor zowel mannen als vrouwen.

In het derde hoofdstuk onderzochten we de dynamische interacties tussen criminaliteit, werk en uitkeringen. Hier controleerden we statistisch voor de bovengenoemde vaardigheden. Het hoofdstuk diende om onderscheid te maken tussen economische en sociologische perspectieven op de relatie tussen werkgelegenheid en criminaliteit. We gebruikten uitkeringen als een identificerend mechanisme door te stellen dat uitkeringen alleen financiële winst geven hetgeen als belangrijk wordt verondersteld in de economische theorieën, en dat uitkeringen niet zorgen voor de sociale banden die belangrijk worden geacht in sociologische theorieën voor het verklaren van criminaliteit. We vonden dat een aanzienlijk deel van de relatie tussen werk en criminaliteit spurious was. De statistische controlevariabelen voor werk en criminaliteit bleken negatief gecorreleerd wat erop wijst dat individuen die gemiddeld een hogere neiging tot crimineel gedrag toonden ook minder vaak werk hadden. In dit hoofdstuk werden geen pogingen gedaan om dit spurious deel van de relatie te interpreteren, maar zoals we hieronder uitleggen kan een deel van deze spurious relatie worden verklaard door verschillen in cognitieve en sociale vaardigheden. In plaats van het interpreteren van de spurious relatie tussen criminaliteit en werk, lag de nadruk in dit hoofdstuk op het duiden van de structurele relaties tussen criminaliteit, werk en uitkeringen. We vonden significante bi-directionele negatieve dynamische structurele effecten tussen werk en criminaliteit. Dit impliceert dat werk toekomstige criminaliteit vermindert en dat criminaliteit de kans op werk vermindert. Enkele verfijningen van het model toonden aan dat slechts “reguliere”
arbeid en niet uitzendwerk in staat was om vermogenscriminaliteit te verminderen, terwijl zowel gewelds- als vermogenscriminaliteit grote negatieve gevolgen bleken te hebben voor de toekomstige kansen op werk. De relatie tussen uitkeringen en criminaliteit was gecomprimeerder. Alleen bijstandsuitkeringen verminderden de kans op vermogenscriminaliteit. Het effect van bijstandsuitkeringen op vermogenscriminaliteit was net zo groot als het effect van werk op criminaliteit. Dit benadrukt de belangrijke rol van geld, en economische theorieën, voor het verklaren van het structurele deel van de relatie tussen werk en vermogensdelen voor achtergestelde jongeren.


Naast deze aanhoudende gevolgen van jeugdige vaardigheden en signalen vonden we vele structurele dynamische interacties tussen de volwassen uitkomstvariabelen. De belangrijkste bevindingen met betrekking tot de structurele relaties tussen criminaliteit, werk en uitkeringen, die in hoofdstuk drie werden gevonden, werden in dit hoofdstuk bevestigd. Daarbovenop vonden we in dit hoofdstuk dat de kans op criminaliteit werd
verminderd door intieme relaties, terwijl drugsgebruik de kans hierop verhoogde. Voor vrouwen bleken intieme relaties meer verklarende kracht te hebben dan voor mannen. Voor vrouwen voorspellen intieme relaties verminderingen in uitkeringen gedurende de adolescentie en ze zijn negatief verbonden met drugsgebruik. Om de vele marginale effecten samen te analyseren maken we ten laatste gebruik van simulatiemethoden. Hypothetische toenames in cognitieve en sociale vaardigheden genereerden grote reducties in criminaliteit. Deze reducties zijn persistent in de zin dat ze criminaliteit verlagen gedurende de gehele adolescentie en volwassen levensduur. Sociale vaardigheden bleken belangrijker bij het verminderen van criminaliteit dan cognitieve vaardigheden. Daarenboven nam door het vergroten van sociale vaardigheden ook de kans werk aanzienlijk toe gedurende de volwassenheid. Voor vrouwen leidde toename van sociale vaardigheden ook tot een significante verlaging van de kans op intieme relaties en drugsgebruik tijdens de adolescentie. Al met al laten de analyses zien dat toenames van sociale vaardigheden leiden tot grote winsten op meerdere adolescenten en volwassen uitkomsten voor zowel mannen en vrouwen.

Uit het laatste hoofdstuk blijkt dat voor achtergestelde gezinnen de relatie tussen werkloosheid en vermogensdelicten veranderd is gedurende de afgelopen eeuw. Tussen 1930 en 1960 vinden we geen significante relatie tussen werkloosheid en vermogensdelicten, terwijl vanaf 1960 tot 2005 er een steeds sterker positief effect van werkloosheid op vermogenscriminaliteit is. Deze bevinding is opmerkelijk als we ons realiseren dat tijdens de grote depressie in de jaren 1930 de werkloosheid zeer hoog was in Nederland. Deze bevinding laat zien dat voorzichtigheid is geboden bij het generaliseren van resultaten uit het verleden naar de toekomst. Vergelijkbare veranderingen in de relatie tussen criminaliteit en intieme relaties zijn te vinden in Beijers et al. (2012).