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

G. Mesters[a,b,d]* , V. van der Geest[a,c] and C. Bijleveld[a,c]

[a] Netherlands Institute for the Study of Crime and Law Enforcement,
[b] Department of Economics and Business, Universitat Pompeu Fabra,
[c] Department of Law, VU University Amsterdam,
[d] Barcelona Graduate School of Economics

February 26, 2015

*Corresponding author: G. Mesters, Universitat Pompeu Fabra, Barcelona GSE & The Netherlands Institute for the Study of Crime and Law Enforcement. Contact address: Universitat Pompeu Fabra, Department of Economics and Business, Ramon Trias Fargas 2527, 08005 Barcelona, Spain, tel. nr. +34 672 124 872, email: geert.mesters@upf.edu

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

Abstract

Objectives: We seek evidence for economic and social mechanisms that aim to explain the relationship between employment and crime. We use the distinctive features of social welfare for identification.

Methods: We consider a sample of disadvantaged males from The Netherlands who are observed between ages 18 and 32 on a monthly time scale. 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.

Results: We find significant negative bi-directional structural effects between employment and property crime. Robustness checks show that only regular employment is able to significantly reduce the offending probability. Further, a significant uni-directional effect is found for the public assistance category of social welfare on property offending.

Conclusions: The results highlight the importance of economic incentives for explaining the relationship between employment and crime for disadvantaged individuals. For these individuals the crime reducing effects from the public assistance category of social welfare are statistically equivalent to those from employment, which suggests the importance of financial gains. Further, the results suggest that stigmatizing effects from offending severely reduce future employment probabilities.

Keywords: dynamic discrete choice, social control, state dependence, reciprocal, unobserved heterogeneity.

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* economic or social mechanisms have more weight in explaining the negative relationship between employment and crime. The economic perspective stresses the trade-off between legal and illegal means for maximizing income (e.g.,

1

Grogger, 1998; Freeman, 1999; Chalfin & Raphael, 2011), whereas the sociological perspective stresses concepts such as social bonds and identity transformations (Laub & Sampson, 2003). Distinguishing between these mechanisms is of vital importance for designing job market programs for offenders, i.e. from a policy perspective. In addition, the 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., Zhang, 1997; Fishback, Johnson, & Kantor, 2010). 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).

In this paper we exploit the distinctive role of social welfare in order to find evidence in favor of economic and social mechanisms that explain the negative association between employment and crime. Heuristically speaking the argument goes as follows. If employment reduces crime and social welfare reduces crime by the same magnitude, we argue that economic gains are an important mechanism. Otherwise, if employment reduces crime and social welfare does not reduce crime we argue that the social aspects of employment have a larger role in reducing crime compared to economic gains. Ultimately, when applying this reasoning to different types of crime (property and violent), employment (regular and temporary) and social welfare (insurance and assistance), the framework enables us to differentiate more precisely between the economic and social mechanisms that aim to explain the relationship between employment and crime.

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, Blokland, & Bijleveld, 2009a, 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 & MacKinnon, 2004, Chap-

---

[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

ter 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 a dynamic model. 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). We provide a detailed discussion for this novel model and its assumptions.

The remainder of this paper is organized as follows. In Section 2 we explain the basics of the welfare system in The Netherlands. We emphasize that the system shares many properties with other Western societies. In Section 3 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 4. The data origins and constructed variables are discussed in Section 5. In Sections 6 and 7 we present the estimation results, while Section 8 provides a general discussion for the results.

2   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 ($825) per month for an 18 year old and €1596 ($1815) 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 employment. Two broad categories of social welfare

---

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.

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 currently have been employed for at least 26 weeks out of the 36 weeks prior to losing ones 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 payments are partially covered by employers and partially by the government. Mental or physical illness must always be confirmed 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 resident 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)[2]. It is important to note that social welfare payments in The Netherlands are never declined because of criminal history. However, when an individual is being incarcerated the eligibility to social welfare stops temporarily.

3 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

---

[2]The other category that is included in the 1.4 million is maternity leave (42.800), which is not included in the current study.

linkages, include economic rational choice theory (Becker, 1968; Ehrlich, 1973; Cook, 1980, 1986), 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 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), Block and Heineke (1975) and Cook (1980, 1986). 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 (2012)).

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 (Zhang, 1997). 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 is 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 coworkers (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 identifying 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 2). If anything, the identity formation caused by criminal behavior will lead to individuals being less reluctant to apply for, or accept, social welfare.

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 social mechanisms that aim to explain the relationship between crime and employment. First, we acknowledge that no single model can distinguish between all the different theoretical mechanisms 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, 1981a, 1981c). 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 asses the importance of stigmatization suggested by 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, economic strain, general strain and social control theories is more complicated since all imply that employment should reduce offending. We summarize the main predictions made by these theories for the effects of employment and social welfare on crime in Table 1. The identifying power comes from including social welfare and differentiating between property and violent crimes. Economic rational choice theory and economic strain theory suggest that social welfare should reduce property offending as it decreases the relative returns from property offending and reduces financial inequality. Social control theory and more general strain theory expect no effect from social welfare on property offending. Further, for violent offending we would expect the same relationships except for that economic theory would not expect violent offending to reduce by means of employment as there is little to trade between employment and violent offending (Kelly, 2000). The different implied patterns of signs and coefficients allows us to distinguish between the

Effect/Theory	Economic Rational	Economic Strain	Social Strain	Social Control
Employment on Property Crime	< 0	< 0	< 0	< 0
Social Welfare on Property Crime	< 0	< 0	$= 0$	$= 0$
Employment on Violent Crime	$= 0$	< 0	< 0	< 0
Social Welfare on Violent Crime	$= 0$	< 0	$= 0$	$= 0$

Table 1: Theoretical implications for the signs of the coefficients for the effects of employment and social welfare on property and violent offending.

different mechanisms that are implied by the theories.

To strengthen our arguments we use different types employment and welfare benefits. In particular, by exploiting differences between regular and temporary employment and insurance policies and public assistance, we make additional arguments for the economic and sociological perspectives.

4 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 (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. We organized this section as follows. We first present the model and its statistical properties. This is followed by a detailed discussion for the features of the model and its underlying assumptions.

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

$$C_{i,t} = \begin{cases} 1 & \text{w.p.} \quad \pi_{C,i,t} \\ 0 & \text{w.p.} \quad 1 - \pi_{C,i,t} \end{cases} , \quad E_{i,t} = \begin{cases} 1 & \text{w.p.} \quad \pi_{E,i,t} \\ 0 & \text{w.p.} \quad 1 - \pi_{E,i,t} \end{cases} , \quad W_{i,t} = \begin{cases} 1 & \text{w.p.} \quad \pi_{W,i,t} \\ 0 & \text{w.p.} \quad 1 - \pi_{W,i,t} \end{cases} ,$$

$$(1)$$

where $C_{i,t} = 1$ when individual i commits an offense in time period t and $C_{i,t} = 0$ otherwise. The outcome $C_{i,t} = 1$ occurs with probability (w.p.) $\pi_{C,i,t}$ and the outcome $C_{i,t} = 0$ with probability $1 - \pi_{C,i,t}$. 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 $i = 1, \dots, N$ and $t = 1, \dots, T$. The probabilities for the offending, employment and welfare outcomes are summarized in $\pi_{i,t} = (\pi_{C,i,t}, \pi_{E,i,t}, \pi_{W,i,t})'$,

9

where $\pi_{C,i,t}$ is the probability for offending, $\pi_{E,i,t}$ is the probability for employment and $\pi_{W,i,t}$ is the probability for social welfare. A similar framework for individual-level decision problems in crime is considered in Durlauf, Navarro, and Rivers (2010).

We model the transformed probabilities $\theta_{i,t} = \log\left[\pi_{i,t}/(1 + \pi_{i,t})\right]$ simultaneously by[3]

$$\theta_{i,t} = \Gamma Y_{i,t-1} + X_{i,t}\beta + \chi_{i,t}, \tag{2}$$

where Γ is the 3×3 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 β and $\chi_{i,t}$ is the 3×1 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

$$\chi_{j,i,t} = \mu_{j,i}f_{j,t}, \qquad \mu_i = (\mu_{C,i}, \mu_{E,i}, \mu_{W,i})', \qquad \mu_i \sim N(\delta, \Sigma_\mu), \tag{3}$$

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 allow to vary over time according to the common factor $f_{j,t}$. The loadings are assumed normally distributed with mean δ and variance matrix Σ_μ. The off-diagonal elements of the matrix Σ_μ 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.

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

$$\delta = \delta_0 + \left(T^{-1}\sum_{t=1}^{T} X_{i,t}\right)\delta_1 + \lambda Y_{i,0}, \tag{4}$$

where δ_0 is the mean vector, δ_1 captures the correlation between the individual-specific effects and the means of the explanatory variables and λ 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

---

[3]Modeling the actual probabilities $\pi_{i,t}$ is difficult as these are restricted to lie between zero and one. The logistic transformation ensures that the transformed probabilities $\theta_{i,t}$ are unrestricted. Note that the 3×1 vector $\theta_{i,t}$ has elements $\log\left[\pi_{C,i,t}/(1 + \pi_{C,i,t})\right]$, $\log\left[\pi_{E,i,t}/(1 + \pi_{E,i,t})\right]$ and $\log\left[\pi_{W,i,t}/(1 + \pi_{W,i,t})\right]$.

nonlinear models in Wooldridge (2005) [4].

The model parameters are summarized in the vector ψ, which contains the unrestricted elements of Γ, β, Σ_μ, δ_0, δ_1 and λ, 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.

Discussion of statistical model

The complete multivariate logistic panel data model is summarized by equations (1), (2), (3) and (4). The coefficients of the matrix Γ are of main interest in our study and we discuss under which assumptions we can estimate them consistently. The diagonal elements of the matrix Γ capture the state dependence in the offending, employment and social welfare outcomes. In other words, they estimate the predictive effect of offending, employment, or receiving social welfare at age $t-1$ on the corresponding transformed probabilities for offending, employment and social welfare at age t. The off-diagonal elements of Γ capture the predictive cross-effects among offending, employment and social welfare. We refer to the parameters in Γ as representing structural effects.

First, regardless of the control variables, we emphasize that we are not estimating contemporaneous effects. For example, we are not estimating the effect of employment at age t on the probability for offending at age t. Instead, the estimates for the parameters in Γ pertain to predictive effects of outcomes at age $t-1$ on the probabilities for the outcomes at age t. Getting consistent estimates of effects for the same age periods requires instrumental variables which are unavailable to us, see the discussion in Section 1. By lagging the outcomes by one monthly period we circumvent the simultaneity problem. Nevertheless, we argue that the predictive structural estimates remain very relevant from a policy perspective. For instance if being employed predicts a significantly lower offending probability for the next month this would remain a useful intervention.

Without additional control variables the approach so far does not remedy the omitted variable problem, which as discussed above was the second source for endogeneity. To explain how the model controls for unobserved heterogeneity first consider a simplified model for only the offending outcome

$$C_{i,t} = \begin{cases} 1 & \text{w.p.} & \pi_{C,i,t} \\ 0 & \text{w.p.} & 1 - \pi_{C,i,t} \end{cases} \qquad \theta_{C,i,t} = \log\left[\pi_{C,i,t}/(1 + \pi_{C,i,t})\right] \tag{5}$$

---

[4]The functional form of this specification can be relaxed to allow for different age-dependent correlations

$$\delta = \delta_0 + \sum_{t=1}^{T} X_{i,t}\delta_t + \lambda Y_{i,0},$$

which is the specification adopted in Chamberlain (1980). Here the coefficients δ_t capture the correlation between the random part of the factor structure and the explanatory variables at each monthly period. We experimented with such extensions and the results did not qualitatively change.

with

$$\theta_{C,i,t} = \Gamma_{11}C_{i,t-1} + X_{i,t}\beta_C + \mu_{C,i}f_{C,t}. \tag{6}$$

The model in equations (5) and (6) represents a uni-variate dynamic binary choice model, see for example Hyslop (1999). The model offers two explanations (state dependence and unobserved heterogeneity) for explaining why offending in period $t-1$ can predict offending in period t. First, state dependence, as captured by $\Gamma_{11}C_{i,t-1}$, suggests that structural features of offending have a direct effect on the offending probability in the next period. Think for example about the financial gains that offending might provide which could be a direct motivation to continue offending. Second, observed and unobserved variables can provide reasons for continuing offending. Unobserved heterogeneity, as captured by $\mu_{C,i}f_{C,t}$, suggests that individuals continue to offend because of underlying unobserved characteristics, such as abilities and preferences. These different explanations for persistence in offending outcomes have extensively been discussed from a theoretical perspective in Nagin and Paternoster (2000).

In other to further clarify the component $\mu_{C,i}f_{C,t}$, suppose that $f_{C,t} = 1$, for all $t = 1, \ldots, T$. The model now reduces to a standard random effects logistic model, where $\mu_{C,i}$ captures the unobserved mean differences for offending, see Baltagi (2005). By allowing $f_{j,t}$ to vary with age, we can allow for the magnitude of the unobserved preferences and abilities to vary with age. It is important to note that if and only if the observed ($X_{i,t}\beta_C$) and unobserved control variables ($\mu_{C,i}f_{C,t}$) adequately capture unobserved heterogeneity the estimate for the state dependence parameter Γ_{11} is consistent. It is possible to check whether this is the case by allowing for more than one common factor $f_{j,t}$ per outcome variable. For example with two factors we would consider $\mu_{1,C,i}f_{1,C,t} + \mu_{2,C,i}f_{2,C,t}$ as the control variable. Now if the variance of $\mu_{i,2}$ becomes zero this would imply that the additional factor is capturing no additional variation. In principal, we could add many factors in this way. However, in our empirical application we found no evidence for multiple factors. A similar univariate model as in (5) and (6) can be formulated for the employment and social welfare outcomes. The arguments for the distinction between state dependence and unobserved heterogeneity are equally valid for these outcomes, albeit with different interpretation.

Now we return to the complete model as summarized by equations (1), (2), (3) and (4). In addition to the univariate model, this model can additionally account for cross-effects among offending, employment and social welfare. For example, following a similar reasoning as above, if employment at age t predicts offending at age $t+1$ this can be due to the predictive structural cross effect from employment on offending or due to correlated unobserved heterogeneity. The structural effect is captured by the parameter Γ_{12}, while correlation between the unobserved components is captured by the correlation between $\mu_{C,i}$ and $\mu_{E,i}$ as modeled by the parameter $\Sigma_{\mu,2,1}$. Recall, that the unobserved random effects $\mu_i = (\mu_{C,i}, \mu_{E,i}, \mu_{W,i})'$ are simultaneously modeled with distribution $\mu_i \sim N(\delta, \Sigma_\mu)$. The off-diagonal elements of Σ_μ capture the spurious correlation between the random effects $(\mu_{C,i}, \mu_{E,i}, \mu_{W,i})'$.

The same distinction between structural cross-effects and spurious cross-effects holds

for all combinations among offending, employment and social welfare. If the estimated off-diagonal elements of Σ_μ are equal to zero we could estimate univariate models for the outcomes. In particular, for offending a sufficient model would be

$$C_{i,t} = \begin{cases} 1 & \text{w.p.} & \pi_{C,i,t} \\ 0 & \text{w.p.} & 1 - \pi_{C,i,t} \end{cases} \qquad \theta_{C,i,t} = \log\left[\pi_{C,i,t}/(1 + \pi_{C,i,t})\right] \qquad (7)$$

with

$$\theta_{C,i,t} = \Gamma_{11}C_{i,t-1} + \Gamma_{12}E_{i,t-1} + \Gamma_{13}W_{i,t-1} + X_{i,t}\beta_C + \mu_{C,i}f_{C,t}. \qquad (8)$$

When the off-diagonal elements of Σ_μ are a priori set to zero while they are really not zero, then the spurious correlation is typically attributed to the structural cross-effects. In equation (8) these effects would be captured by Γ_{12} and Γ_{13}, but they should have been attributed to $\Sigma_{\mu,2,1}$ and $\Sigma_{\mu,3,1}$. Typically, under the assumption that the true spurious correlations between offending and employment and offending and social welfare are negative, the cross dependence parameters would be overestimated.

The distinction between spurious cross-effects and structural cross-effects is important for understanding the dynamic relationships among offending, employment and social welfare. For example, if the parameter Γ_{12}, which captures the structural effect of employment on offending is not significant but $\Sigma_{\mu,2,1}$ is significant and negative, than this would imply that individuals who select with higher probability into employment at the same time select with lower probability into offending. In this scenario there would be no advantage of increasing employment opportunities since the negative correlation is entirely spurious. In our empirical study we show the results from the trivariate model and the restricted uni-variate models to demonstrate the differences.

5 Data

This section outlines the observed data that we use for estimating the parameters of the discrete choice model in Section 4. The sample that we consider consists of $N = 270$ males who where 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 where 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, van der Geest, & Bijleveld, 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 pertaining to a disadvantaged high risk sample (van der Geest, Blokland, & Bijleveld, 2009b). Typically, disadvantaged individuals are responsible for a large portion of the total offenses. 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.

5.1 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 counterfeiting, 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.

5.2 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 2).

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 constract. 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 imply 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. At least once over the observational period 34.4% of the individuals received unemployment insurance benefits, 13.3% received disability benefits and 35.5% received public assistance.

5.3   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 for 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 variables 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 they got 0.56 children.

A special situation occurs when an individual is incarcerated. In this situation, which occurs 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[5].

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.

5.4   Summary statistics

Summary statistics for offending, employment, social welfare and the control variables are given in Table 2. 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 were 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% obtained 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 considered very high given the young ages.

In Figure 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

---

[5]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.

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. We find that property offending is negatively correlated with all categories of employment and social welfare. The correlation between property offending and total employment 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 policies are not correlated with violent offending. We emphasize that Table 3 shows unconditional correlations. As such when a certain outcome occurs in relatively few periods while the other has high variability over the entire observational period the unconditional correlation becomes low in magnitude. This explains for example the low correlation between disability insurance, which is observed in few periods, and employment.

6   Trivariate model results

In this section we use the econometric model of Section 4 to investigate whether the correlations shown in Table 3 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 5. In Section 7 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

Category	Variable	Mean	SD
Property offending	$C_{i,t}^p$	0.026	0.160
Violent offending	$C_{i,t}^v$	0.001	0.094
Employment	$E_{i,t}$	0.314	0.464
Employment regular	$E_{i,t}$	0.232	0.422
Employment temporary	$E_{i,t}$	0.087	0.281
Total welfare	$W^{i,t}$	0.100	0.300
Unemployment insurance	$W_{i,t}^u$	0.023	0.151
Disability insurance	$W_{i,t}^d$	0.006	0.076
Public assistance	$W_{i,t}^d$	0.071	0.257
Marriage	$X_{i,t,1}$	0.070	0.255
Divorce	$X_{i,t,2}$	0.025	0.156
Parenthood	$X_{i,t,3}$	0.204	0.553
Detention	$X_{i,t,4}$	0.095	0.293

Table 2: Summary statistics for all variables for the full sample of individuals ($N = 270$) and age groups ($T = 168$).

Variable	$C_{i,t}^p$	$C_{i,t}^v$	$E_{i,t}$	$E_{i,t}^r$	$E_{i,t}^t$	$W_{i,t}$	$W_{i,t}^u$	$W_{i,t}^d$	$W_{i,t}^p$
$C_{i,t}^p$	1.000								
$C_{i,t}^v$	0.118	1.000							
$E_{i,t}$	-0.074	-0.024	1.000						
$E_{i,t}^r$	-0.075	-0.026	0.811	1.000					
$E_{i,t}^t$	-0.011	-0.000	0.455	-0.135	1.000				
$W_{i,t}$	-0.026	0.004	-0.182	-0.156	-0.070	1.000			
$W_{i,t}^u$	-0.013	0.000	-0.054	-0.051	-0.011	0.466	1.000		
$W_{i,t}^d$	-0.011	-0.001	-0.020	-0.020	-0.005	0.228	0.025	1.000	
$W_{i,t}^p$	-0.019	0.005	-0.176	-0.148	-0.074	0.831	-0.040	-0.019	1.000

Table 3: Correlations among the dependent variables for the full sample of individuals ($N = 270$) and age groups ($T = 168$).
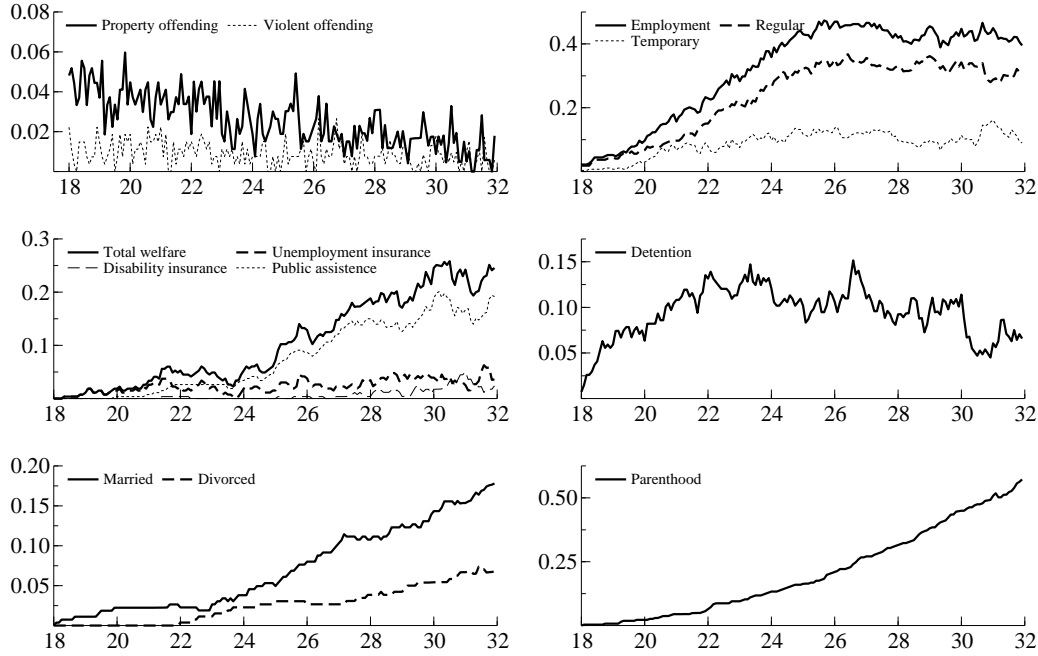
Figure 1: Averages over the full sample of individuals ($N = 270$) per age-period ($T = 168$).

equations (1), (2) and (3), are given in Table 4. 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 Γ and the spurious parameters β, δ and Σ_μ. 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 can follow as social welfare

19

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 their expected signs, but since these coefficients are potentially inconsistently estimated we do not interpret their magnitudes. The most important spurious coefficients are also given in Table 4. 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 2 we show the estimated factors which are modeled using cubic spline functions. We have scaled the factors with the means δ of the unobserved heterogeneous effects μ_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 μ_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 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 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
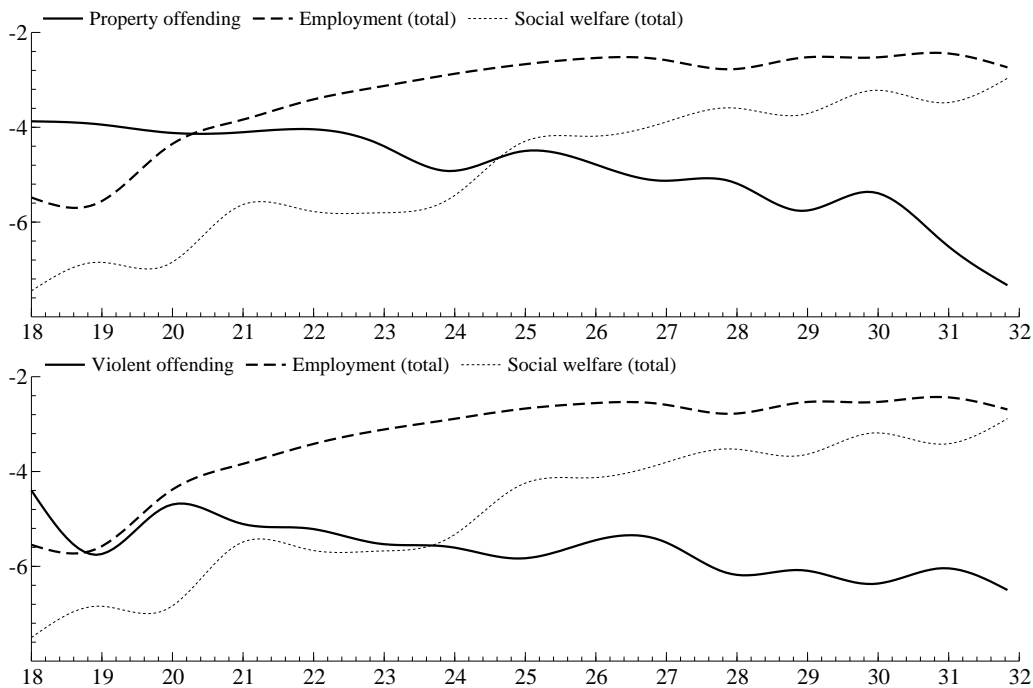
Figure 2: Smooth factors for offending, employment and social welfare.

important factor is that our sample is relatively small and violent offenses are rare, even for the high risk sample under consideration.

6.1 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 5 we show the structural parameters that are obtained by restricting the model to be uni-directional and by considering the model with time-invariant individual-specific effects. The former reduces the model to the standard random effects model with lagged explanatory variables (Baltagi, 2005), while the latter considers equation (3) for $f_{j,t} = 1$ for all j and t. We omitted the parameters for the control variables from Table 5.

The uni-directional models tend to overstate the cross-outcome structural effects. This confirms the motivation for the trivariate model as discussed in Section 4. 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

| | | **Property offending** | | **Violent offending** | |
Variable	Parameter	Estimate	SD	Estimate	SD
State dependence offending	$\Gamma_{1,1}$	1.280*	0.097	0.892*	0.263
Employment on offending	$\Gamma_{1,2}$	-0.424*	0.123	-0.040	0.173
Social welfare on offending	$\Gamma_{1,3}$	-0.151	0.197	0.097	0.225
State dependence employment	$\Gamma_{2,2}$	5.627*	0.061	5.630*	0.061
Offending on employment	$\Gamma_{2,1}$	-0.902*	0.234	-1.151*	0.335
Social welfare on employment	$\Gamma_{2,3}$	-1.852*	0.125	-1.843*	0.125
State dependence social welfare	$\Gamma_{3,3}$	6.354*	0.107	6.383*	0.109
Offending on social welfare	$\Gamma_{3,1}$	0.162	0.366	0.664	0.464
Employment on social welfare	$\Gamma_{3,2}$	-1.416*	0.130	-1.390*	0.131
Marriage on offending	$\beta_{1,1}$	-0.214	0.346	-0.585	0.363
Divorce on offending	$\beta_{1,2}$	0.302	0.598	-0.548	0.544
Parenthood on offending	$\beta_{1,3}$	-0.564*	0.222	0.287	0.167
Incarceration on offending	$\beta_{1,4}$	0.004	0.004	0.001	0.005
Marriage on employment	$\beta_{2,1}$	0.280*	0.139	0.291*	0.140
Divorce on employment	$\beta_{2,2}$	0.439*	0.222	0.459*	0.222
Parenthood on employment	$\beta_{2,3}$	-0.061	0.071	-0.050	0.072
Incarceration on employment	$\beta_{2,4}$	-0.008*	0.003	-0.010*	0.003
Marriage on social welfare	$\beta_{3,1}$	-0.260	0.206	-0.260	0.206
Divorce on social welfare	$\beta_{3,2}$	0.369	0.317	0.337	0.315
Parenthood on social welfare	$\beta_{3,3}$	-0.201	0.103	-0.195	0.103
Incarceration on social welfare	$\beta_{3,4}$	-0.001	0.004	-0.006	0.004
Mean offending	$\delta_{0,1}$	-3.849*	0.147	-5.214*	0.226
Mean employment	$\delta_{0,2}$	-5.389*	0.177	-5.397*	0.177
Mean social welfare	$\delta_{0,3}$	-7.729*	0.396	-7.584*	0.385
Variance offending	$\Sigma_{\mu,1,1}$	1.787*	0.268	1.857*	0.383
Covariance offending employment	$\Sigma_{\mu,2,1}$	-0.462*	0.193	-0.299	0.228
Covariance offending social welfare	$\Sigma_{\mu,3,1}$	-0.492	0.286	0.369	0.326
Variance employment	$\Sigma_{\mu,2,2}$	2.531*	0.403	2.549*	0.404
Covariance employment social welfare	$\Sigma_{\mu,3,2}$	1.184*	0.347	1.042*	0.336
Variance social welfare	$\Sigma_{\mu,3,3}$	3.835*	0.832	3.480*	0.784

Table 4: 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.

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 2).

Further, one might wonder how robust our findings are with respect to other omitted variables, such as drug use, mental illness or homelessness. As we do not have information for these variables we cannot provide a definite answer for this issue. However, we can investigate the changes in the structural parameters when we omit the observed control variables for marriage, divorce, parenthood and incarceration. If the factor structures that aim to control for unobserved variables are sufficient for capturing the omitted effects from these variables we can gain some more confidence in our results.

In Table 6 we show the results from estimating the model without observed control variables. We find that the structural parameters Γ do not change in any noticeable way. Instead, we do observe large changes in the spurious correlation coefficients as captured by the off-diagonal elements of Σ_μ. In particular, the spurious correlation between employment and property offending has increased in magnitude from -0.462 to -0.599 and the spurious correlation between social welfare and property offending changed from -0.492 to -0.352. This indicates that the correlations that were previously attributed to marriage, divorce, parenthood and incarceration are now picked up by the latent factor structure, leaving the structural parameters of interest unchanged.

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.

7   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 4 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 6. This allows us to compare the results for the structural parameters to those given in Table 4.

The parameter estimates for the regular and temporary employment variables are given in Table 7. 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 4 (-0.736 vs -0.424). Recall that regular employment refers to employment that is registered on the payroll of the

	Parameter	**Property offending** Estimate	SD	**Violent offending** Estimate	SD
Uni-directional random effects models					
State dependence offending	$\Gamma_{1,1}$	1.271^*	0.097	0.833^*	0.265
Employment on offending	$\Gamma_{1,2}$	-0.480^*	0.122	-0.111	0.166
Social welfare on offending	$\Gamma_{1,3}$	-0.247	0.195	0.129	0.219
State dependence employment	$\Gamma_{2,2}$	5.595^*	0.193	5.597^*	0.193
Offending on employment	$\Gamma_{2,1}$	-0.960	0.738	-1.209	1.058
Social welfare on employment	$\Gamma_{2,3}$	-1.772^*	0.395	-1.768^*	0.395
State dependence social welfare	$\Gamma_{3,3}$	6.352^*	0.343	6.350^*	0.343
Offending on social welfare	$\Gamma_{3,1}$	0.079	1.151	0.721	1.479
Employment on social welfare	$\Gamma_{3,2}$	-1.287^*	0.404	-1.293^*	0.403
Time-invariant individual-specific effects					
State dependence offending	$\Gamma_{1,1}$	1.300^*	0.097	0.840^*	0.263
Employment on offending	$\Gamma_{1,2}$	-0.676^*	0.116	-0.060	0.163
Social welfare on offending	$\Gamma_{1,3}$	-0.568^*	0.180	-0.086	0.215
State dependence employment	$\Gamma_{2,2}$	6.011^*	0.060	6.006^*	0.060
Offending on employment	$\Gamma_{2,1}$	-1.020^*	0.234	-1.203^*	0.330
Social welfare on employment	$\Gamma_{2,3}$	-1.513^*	0.133	-1.499^*	0.133
State dependence social welfare	$\Gamma_{3,3}$	6.918^*	0.105	6.951^*	0.106
Offending on social welfare	$\Gamma_{3,1}$	-0.076	0.355	0.493	0.461
Employment on social welfare	$\Gamma_{3,2}$	-0.815^*	0.133	-0.801^*	0.133

Table 5: Structural parameter estimates for the uni-directional models and the trivariate logistic panel data model with time-invariant unobserved heterogeneity. 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.

Variable	Parameter	Property offending		Violent offending	
		Estimate	SD	Estimate	SD
State dependence offending	$\Gamma_{1,1}$	1.277*	0.096	0.893*	0.261
Employment on offending	$\Gamma_{1,2}$	-0.423*	0.122	-0.050	0.171
Social welfare on offending	$\Gamma_{1,3}$	-0.169	0.196	0.135	0.222
State dependence employment	$\Gamma_{2,2}$	5.642*	0.060	5.645*	0.061
Offending on employment	$\Gamma_{2,1}$	-0.888*	0.234	-1.152*	0.333
Social welfare on employment	$\Gamma_{2,3}$	-1.860*	0.125	-1.857*	0.125
State dependence social welfare	$\Gamma_{3,3}$	6.357*	0.106	6.384*	0.107
Offending on social welfare	$\Gamma_{3,1}$	0.169	0.367	0.640	0.468
Employment on social welfare	$\Gamma_{3,2}$	-1.414*	0.130	-1.391*	0.130
Mean offending	$\delta_{0,1}$	-3.856*	0.148	-5.230*	0.217
Mean employment	$\delta_{0,2}$	-5.421*	0.180	-5.440*	0.181
Mean social welfare	$\delta_{0,3}$	-7.715*	0.388	-7.511*	0.366
Variance offending	$\Sigma_{\mu,1,1}$	1.844*	0.260	1.852*	0.305
Covariance offending employment	$\Sigma_{\mu,2,1}$	-0.599*	0.190	-0.358	0.210
Covariance offending social welfare	$\Sigma_{\mu,3,1}$	-0.352	0.270	0.270	0.257
Variance employment	$\Sigma_{\mu,2,2}$	2.693*	0.429	2.737*	0.438
Covariance employment social welfare	$\Sigma_{\mu,3,2}$	1.231*	0.363	1.196*	0.349
Variance social welfare	$\Sigma_{\mu,3,3}$	3.805*	0.803	3.288*	0.690

Table 6: Parameters estimates for the trivariate logistic panel data model without observed control variables and 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.

employer. The effect of temporary employment (via a 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 8. 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.

8   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 paper was to distinguish between economic and social mechanisms that aim to explain the relationship between employment and crime. The role of social welfare was used as an identifying mechanism. The economic mechanisms were motivated by economic rational choice and economic strain theories, whereas social control theory and general strain theory provided the social mechanisms that we considered. Further, we used labeling theory to predict a negative structural effect of offending on employment.

A distinction between economic and social mechanisms 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, social mechanisms 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.

	Parameter	**Property offending**		**Violent offending**	
		Estimate	SD	Estimate	SD
Regular employment					
State dependence offending	$\Gamma_{1,1}$	1.283*	0.096	0.901*	0.262
Employment on offending	$\Gamma_{1,2}$	-0.763*	0.188	-0.197	0.217
Social welfare on offending	$\Gamma_{1,3}$	-0.151	0.194	0.033	0.221
State dependence employment	$\Gamma_{2,2}$	6.282*	0.076	6.292*	0.076
Offending on employment	$\Gamma_{2,1}$	-0.617	0.356	-0.902	0.471
Social welfare on employment	$\Gamma_{2,3}$	-1.901*	0.174	-1.901*	0.174
State dependence social welfare	$\Gamma_{3,3}$	6.473*	0.107	6.504*	0.108
Offending on social welfare	$\Gamma_{3,1}$	0.125	0.367	0.659	0.469
Employment on social welfare	$\Gamma_{3,2}$	-1.228*	0.155	-1.210*	0.154
Temporary employment					
State dependence offending	$\Gamma_{1,1}$	1.294*	0.097	0.907*	0.263
Employment on offending	$\Gamma_{1,2}$	-0.137	0.141	0.027	0.205
Social welfare on offending	$\Gamma_{1,3}$	-0.063	0.195	0.129	0.220
State dependence employment	$\Gamma_{2,2}$	5.257*	0.070	5.253*	0.070
Offending on employment	$\Gamma_{2,1}$	-0.741*	0.247	-1.027*	0.393
Social welfare on employment	$\Gamma_{2,3}$	-1.145*	0.148	-1.157*	0.147
State dependence social welfare	$\Gamma_{3,3}$	6.656*	0.105	6.675*	0.106
Offending on social welfare	$\Gamma_{3,1}$	0.252	0.373	0.675	0.465
Employment on social welfare	$\Gamma_{3,2}$	-1.053*	0.197	-1.049*	0.197

Table 7: 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.

	Parameter	**Property offending**		**Violent offending**	
		Estimate	SD	Estimate	SD
Unemployment insurance					
State dependence offending	$\Gamma_{1,1}$	1.286*	0.097	0.889*	0.262
Employment on offending	$\Gamma_{1,2}$	-0.401*	0.121	-0.072	0.167
Social welfare on offending	$\Gamma_{1,3}$	0.405	0.312	0.433	0.368
State dependence employment	$\Gamma_{2,2}$	5.751*	0.060	5.753*	0.060
Offending on employment	$\Gamma_{2,1}$	-0.871*	0.234	-1.128*	0.330
Social welfare on employment	$\Gamma_{2,3}$	-1.347*	0.180	-1.361*	0.180
State dependence social welfare	$\Gamma_{3,3}$	6.153*	0.144	6.183*	0.146
Offending on social welfare	$\Gamma_{3,1}$	0.043	0.503	0.076	0.740
Employment on social welfare	$\Gamma_{3,2}$	-0.975*	0.152	-0.977*	0.152
Disability insurance					
State dependence offending	$\Gamma_{1,1}$	1.287*	0.097	0.895*	0.262
Employment on offending	$\Gamma_{1,2}$	-0.423*	0.121	-0.097	0.166
Social welfare on offending	$\Gamma_{1,3}$	-0.321	1.056	0.788	0.572
State dependence employment	$\Gamma_{2,2}$	5.777*	0.060	5.775*	0.060
Offending on employment	$\Gamma_{2,1}$	-0.904*	0.235	-1.135*	0.330
Social welfare on employment	$\Gamma_{2,3}$	-1.317*	0.349	-1.360*	0.350
State dependence social welfare	$\Gamma_{3,3}$	5.907*	0.313	6.143*	0.365
Offending on social welfare	$\Gamma_{3,1}$	-2.519	2.518	-1.266	1.980
Employment on social welfare	$\Gamma_{3,2}$	-0.712*	0.276	-0.589*	0.276
Public Assistance					
State dependence offending	$\Gamma_{1,1}$	1.283*	0.097	0.974*	0.261
Employment on offending	$\Gamma_{1,2}$	-0.448*	0.122	-0.131	0.169
Social welfare on offending	$\Gamma_{1,3}$	-0.490*	0.238	-0.395	0.256
State dependence employment	$\Gamma_{2,2}$	5.670*	0.060	5.674*	0.060
Offending on employment	$\Gamma_{2,1}$	-0.939*	0.235	-1.129*	0.335
Social welfare on employment	$\Gamma_{2,3}$	-2.260*	0.212	-2.193*	0.213
State dependence social welfare	$\Gamma_{3,3}$	7.764*	0.213	7.823*	0.214
Offending on social welfare	$\Gamma_{3,1}$	0.327	0.575	1.189*	0.578
Employment on social welfare	$\Gamma_{3,2}$	-1.753*	0.269	-1.650*	0.271

Table 8: Structural parameter estimates for the trivariate logistic panel data model for different types of social welfare: unemployment insurance, disability insurance and public benefits. 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.

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 social mechanisms 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 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, has 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 is not the only contributing mechanism to the negative correlation between employment and crime, but that the mechanisms implied by 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 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 social and economic mechanisms 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 for 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 mechanisms 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 en 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 4. The methodology is discussed for a more general panel data model in Mesters and Koopman (2014). All parameters are placed in the vector ψ, which includes the unrestricted elements of Γ, β, δ_0, δ_1, λ, Σ_v and the knots of the splines.

We summarize the vector of dependent variables for individual i in time period t by $Y_{i,t} = (Ci, t, E_{i,t}, W_{i,t})'$, which is thus a 3×1 vector of binary variables. The loglikelihood for the observations $Y = \{Y_{i,t}\}_{i=1,\ldots,N\ t=1,\ldots,T}$ 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 μ_i, defined in (3), we can express the joint density as a high dimensional integral as follows

$$p(Y;\psi) = \int_\mu p(Y,\mu;\psi)\ \mathrm{d}\mu = \int_\mu p(Y|\mu;\psi)p(\mu;\psi)\ \mathrm{d}\mu, \qquad (9)$$

where $\mu = \{\mu_i\}_{i=1,\ldots,N}$ and $p(\mu;\psi)$ is defined in equation (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 $\theta_{i,t} = (\theta_{C,i,t}, \theta_{E,i,t}, \theta_{W,i,t})$ is given in (2) and $Y_{j,i,t}$ corresponds to the outcome variables $Y_{C,i,t} = Ci, t$, $Y_{E,i,t} = E_{i,t}$ 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 (9). Instead we follow the conventional literature and solve the integral using Monte Carlo methods. We refer to Cappé, Moulines, and Rydén (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 μ 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 $\mu^{(s)}$ denotes the sth 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 μ 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

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

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

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

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

$$z_i = \mu_i + u_i, \qquad u_i \sim NID(0, D_i),$$

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

Algorithm

(i) Initialize $\mu = \mu^*$;

(ii) Given μ^*; 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 $i = 1, \ldots, N$;

(iii) Update μ^* by computing $\mathrm{E}_g(\mu|z)$ based on $z_i = \mu_i + u_i$ and $u_i \sim NID(0, D_i)$;

(iv) 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 μ_i given in (3). Using these samples we construct the Monte Carlo likelihood. The resulting likelihood estimate $\hat{p}(y; \psi)$ is optimized with respect to parameters ψ 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 $\bar{f}_{R(l),t}$, 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 $\bar{f}_{R(l),t}$ 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.

References

Agnew, R. (1992). Foundation for a General Strain Theory of Crime and Delinquency. *Criminology*, *30*, 47-87.

Ahn, S. G., Lee, Y. H., & Schmidt, P. (2013). Panel data models with multiple time-varying individual effects. *Journal of Econometrics*, *174*, 1-14.

Alessie, R., Hochguertel, S., & van Soest, A. (2004). Ownership of Stocks and Mutual Funds: A Panel Data Analysis. *The Review of Economics and Statistics*, *86*, 783-796.

Apel, R., & Sweeten, G. (2010). The impact of incarceration on employment during the transition to adulthood. *Social Problems*, *57*, 448-479.

Bai, J., & Ng, S. (2009). Panel Data Models With Interactive Fixed Effects. *Econometrica*, *77*, 1229-1279.

Baltagi, B. H. (2005). *Econometric analysis of panel data.* New York: John Wiley & Sons.

Becker, G. S. (1968). Crime and punishment: An economic analysis. *Journal of Political Economy*, *78*, 169-217.

Becker, G. S., & Mulligan, C. B. (1997). The Endogenous Determination of Time Preferences. *The Quarterly Journal of Economics*, *112*, 729-758.

Bersani, B. E., Laub, J. H., & Nieuwbeerta, P. (2009). Marriage and desistance from crime in the netherlands: Do gender and socio-historical context matter? *Journal of Quantitative Criminology*, *25*, 3-24.

Block, M., & Heineke, J. (1975). A Labour Theoretical Analysis of Criminal Choice. *American Economic Review*, *65*, 314-325.

Cappé, O., Moulines, E., & Rydén. (2005). *Inference in Hidden Markov Models.* New York: Springer Verlag.

CBS. (2010). Centraal bureau voor de statistiek: Standaardclassificatie misdrijven. (from http://www.cbs.nl)

CBS. (2012). Centraal bureau voor de statistiek: Crime rates database, The Netherlands. (from http://statline.cbs.nl)

Chalfin, A. J., & Raphael, S. (2011). Work and Crime. In M. Tonry (Ed.), *Oxford Handbook of Crime and Criminal Justice.* Oxford: Oxford University Press.

Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *The Review of Economic Studies*, *47*, 225-238.

Cohen, M. A. (1998). The Monetary Value of Saving a High-Risk Youth. *Journal of Quantitative Criminology*, *14*, 5-33.

Cook, P. J. (1980). Research in criminal deterrence: laying the groundwork for the second decade. In N. Morris (Ed.), *Crime and justice: an annual review of research.* Chicago: University of Chicago Press.

Cook, P. J. (1986). The demand and supply of criminal opportunities. In M. Tonry & N. Morris (Eds.), *Crime and justice: an annual review of research.* Chicago: University of Chicago Press.

Davidson, R., & MacKinnon, J. G. (2004). *Econometric Theory and Methods.* New York: Oxford University Press.

de Mooij, R. A. (2006). Reinventing the Welfare State. (CPB, Netherlands Bureau for Economic Policy Analysis)

Dobkin, C., & Puller, S. (2007). The Effects of Government Transfers on Monthly Cycles in Drugs Abuse, Hospitalization, and Mortality. *Journal of Public Economics*, *91*, 2137-2157.

Durbin, J., & Koopman, S. J. (2012). *Time Series Analysis by State Space Methods; 2nd edition.* Oxford: Oxford University Press.

Durlauf, S. N., Navarro, S., & Rivers, D. A. (2010). Understanding aggregate crime regressions. *Journal of Econometrics*, *158*, 306-317.

Ehrlich, I. (1973). Participation in illegal activities: A theoretical and empirical investigation. *Journal of Political Economy*, *81*, 521-565.

Fishback, P. V., Johnson, R. S., & Kantor, S. (2010). Striking at the Roots of Crime: The Impact of Welfare Spending on Crime during the Great Depression. *Journal of Law and Economics*, *53*, 715-740.

Foley, C. F. (2011). Welfare Payments and Crime. *The Review of Economics and Statistics*, *93*, 97-112.

Freeman, R. B. (1999). The Economics of Crime. In O. Ashenfelter & D. E. Card (Eds.), *Handbook of Labor Economics, vol. 3c* (p. 3529-3571). Amsterdam, North-Holland: Elsevier Science.

Goodman, P. (1956). *Growing Up Absurd.* New York: Random House.

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

Gould, E. D., Weinberg, B. A., & Mustard, D. B. (2002). Crime Rates and Local Labor Market Opportunities in the United States: 1977-1997. *The Review of Economics and Statistics*, *84*, 45-61.

Grogger, J. (1995). The Effect of Arrest on the Employment and Earnings of Young Men. *Quarterly Journal of Economics*, *110*, 51-71.

Grogger, J. (1998). Market Wages and Youth Crime. *Journal of Labor Economics*, *16*, 756-791.

Heckman, J. J. (1981a). Heterogeneity and State Dependence. In S. Rosen (Ed.), *Studies of Labor Markets* (p. 91-140). Chicago: The National Bureau of Economic Research: The University of Chicago Press.

Heckman, J. J. (1981b). The Initial Parameters Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In C. F. Manski & D. L. McFadden (Eds.), *Structural Analysis of Discrete Data and Econometric Applications* (p. 179-195). Cambridge: The MIT Press.

Heckman, J. J. (1981c). Statistical Models for Discrete Panel Data. In C. F. Manski & D. L. McFadden (Eds.), *Structural Analysis of Discrete Data and Econometric Applications* (p. 113-177). Cambridge: The MIT Press.

Hirschi, T. (1969). *Causes of Delinquency.* Berkeley, CA: University of California Press.

Hyslop, D. R. (1999). State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women. *Econometrica*, *67*, 1255-1294.

Jungbacker, B., & Koopman, S. J. (2007). Monte Carlo Estimation for Nonlinear Non-Gaussian State Space Models. *Biometrika*, *94*, 827-839.

Jungbacker, B., Koopman, S. J., & van der Wel, M. (2014). Smooth Dynamic Factor Analysis with Application to the U.S. Term Structure of Interest Rates. *Journal of Applied Econometrics*, *29*, 65-90.

Kelly, M. (2000). Inequality and Crime. *The Review of Economics and Statistics*, *82*, 530-539.

Lageson, S., & Uggen, C. (2013). How Work Affects Crime - And Crime Affects Work - Over The Life Course. In C. L. Gibson & M. D. Krohn (Eds.), *Handbook of Life-Course Criminology: Emerging Trends and Directions for Future Research.* New York: Springer.

Laub, J. H., & Sampson, R. J. (2003). *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70.* Cambridge, MA: Harvard University Press.

Lemert, E. (1967). *Human deviance, social problems, and social control.* Englewood Cliffs: Prentice-Hall.

Levitt, S. D., & Venkatesh, S. A. (2001). An Economic Analysis of a Drug-Selling Gang's Finances. *Quarterly Journal of Economics*, *39*, 755-789.

Lin, M. J. (2008). Does Uemployment Increase Crime? *Journal of Human Resources*, *43*, 413-436.

Machin, S., & Meghir, C. (2004). Crime and Economic Incentives. *Journal of Human Resources*, *39*, 958-979.

Maruna, S. (2001). *Making Good: How Ex-Convicts Reform and Rebuild Their Lives.* Washington, DC: American Psychological Association Books.

Merton, R. K. (1938). Social structure and anomie. *American Sociological Review*, *3*, 672-682.

Mesters, G., & Koopman, S. J. (2014). Generalized Dynamic Panel Data Models with Random Effects for Cross-Section and Time. *Journal of Econometrics*, *180*, 127–140.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, *46*, 69-85.

Mustard, D. B. (2010). How Do Labor Markets Affect Crime? New Evidence on an Old Puzzle. In B. L. Benson & P. R. Zimmerman (Eds.), *Handbook on the Economics of Crime.* New York: Edward Elgar Publishing.

Nagin, D., & Paternoster, R. (2000). Population Heterogeneity and State Dependence : Future Research. *Journal of Quantitative Criminology*, *16*, 117-145.

Nocedal, J., & Wright, S. J. (1999). *Numerical optimization.* New York: Springer Verlag.

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

Pager, D. (2007). *Marked: race, crime and findings work in an era of mass incarceration.* Chicago, IL: University of Chicago Press.

Pager, D., Western, B., & Bonikowski, B. (2009). Discrimination in a low-wage market: a field experiment. *American Sociological Review*, *74*, 777-799.

Pesaran, H. M. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, *74*, 967-1012.

Phelps, E. S. (1994). Low-Wage Employment Subsidies versus the Welfare State. *American Economic Review, Papers and Proceedings*, *84*, 54-58.

Poirier, D. J. (1976). *The econometrics of structural change: with special emphasis on spline functions.* Amsterdam: North-Holland.

Raphael, S. (2011). Improving Employment Prospects for Former Prison Inmates: Challenges and Policy. In P. J. Cook, J. Ludwig, & J. McCrary (Eds.), *Controlling Crime: Strategies and Tradeoffs.* Chicago, Il: University of Chicago Press.

Raphael, S., & Winter-Ebmer, R. (2001). Identifying the Effect of Unemployment on Crime. *Journal of Law and Economics*, *44*, 259-283.

Ripley, B. D. (1987). *Stochastic Simulation.* New York: John Wiley & Sons.

Sampson, R. J., & Laub, J. H. (1993). *Crime in the making: pathways and turning points through life.* Cambridge, MA: Harvard University Press.

Sampson, R. J., & Laub, J. H. (2005). A life-course view of the development of crime. *ANNALS of the American Academy of Political and Social Science*, *602*, 12-45.

Shanahan, M. J. (2000). Pathways to adulthood in changing societies: variability and mechanisms in life course perspective. *Annual Review of Sociology*, *26*, 667-692.

Shapiro, J. (2005). Is There a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle. *Journal of Public Economics*, *89*, 303-325.

Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2007). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: Evidence for a dual systems model. *Developmental Psychology*, *44*, 1764-1777.

Stolzenberg, L. J., & D'Alessio, S. J. (2007). The effect of divorce on domestic crime. *Crime & Delinquency*, *53*, 281-302.

Tannenbaum, F. (1938). *Crime and community.* London and New York: Columbia University Press.

Thornberry, T. P., & Christensen, R. L. (1984). Unemployment and CriminalInvolvement: An Investigation of Reciprocal Causal Structures. *American Sociological Review*, *49*, 398-411.

van der Geest, V., Blokland, A., & Bijleveld, C. (2009a). Delinquent development in a sample of high-risk youth: Shape, content and predictors of delinquent trajectories from age 12 to 32. *Journal of Research in Crime and Delinquency*, *46*, 111–143.

van der Geest, V., Blokland, A., & Bijleveld, C. (2009b). Delinquent development in a sample of high-risk youth: Shape, content and predictors of delinquent trajectories from age 21 to 32. *Journal of Research in Crime and Delinquency*, *46*, 111-143.

van der Geest, V., Blokland, A., & Bijleveld, C. C. J. H. (2011). The effects of employment on longitudinal trajectories of offending: A follow up in high risk youths from ages 18 - 32. *Criminology*, *49*, 1195-1234.

van Oorschot, W. (2006). The Dutch welfare state: recent trends and challenges in historical perspective. *European Journal of Social Security*, *8*, 57-76.

Wijkman, M., van der Geest, V., & Bijleveld, C. (2006). Delinquentie van op civielrechtelijke

en strafrechtelijke titel behandelde jongeren (delinquency after treatment in a juvenile justice institution under civil or criminal law). *Tijdschrift voor Familie- en Jeugdrecht*, *1*, 18–27.

Wilkinson, R., & Pickett, K. (2009). *The spirit level: Why more equal societies almost always do better*. London: Allen Lane.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, *20*, 39-54.

Zhang, J. (1997). The Effect of Welfare Programs on Criminal Behavior: A Theoretical and Empirical Analysis. *Economic Inquiry*, *35*, 120–137.