Barred from employment? A study of labor market prospects before and after imprisonment

Ramakers, A.A.T.

4 Imprisonment length and employment prospects■

Abstract

This study considers the relationship between imprisonment length and employment outcomes. The data come from a unique prospective, longitudinal study of Dutch pretrial detainees (n = 702). All subjects thus experience prison confinement of varying lengths, although the durations are relatively short (mean = 3.8 months; median = 3.1 months). This contrasts with prior research that was limited to the study of American prison sentences spanning an average of 2 years. These data thus fill a gap in the empirical base concerning short-term confinement, which is the norm in the United States (e.g., jail incarceration) and other Western countries. Using a comprehensive array of pre-prison covariates, a propensity score methodology is used to examine the dose-response relationship between imprisonment length and a variety of employment outcomes. The results indicate that, among prison spells less than 6 months in duration, longer confinement is largely uncorrelated with employment. In contrast, among spells in excess of 6 months, longer imprisonment length seems to worsen employment prospects.

Keywords: imprisonment length, employment, propensity score methodology, the Netherlands.

■ A slightly different version is accepted for publication as: Ramakers, A.A.T., Apel, R., Nieuwbeerta, P., Dirkzwager, A., & Van Wilsem, J.A. (forthcoming). Imprisonment length and post-prison employment prospects in the Netherlands. *Criminology.* In production.

4.1 Introduction

In recent years, much scholarly attention has been focused on the social disabilities caused by mass imprisonment (Dumont, Brockmann, Dickman, Alexander, & Rich, 2012; Raphael, 2011; Wakefield & Uggen, 2010; Western, 2006). The alarming scale of incarceration has brought issues of prisoner reentry to the fore (Petersilia, 2003; Visher & Travis, 2003), as prisons experienced three full decades of uninterrupted growth. This prison growth persisted through the most pronounced crime decline of the modern era, and only recently has it begun to slow and even stabilize. Although the United States stands out for its unbridled enthusiasm toward the use of incarceration as a solution to the crime problem, it is hardly unique in its trend of growing punitiveness. Indeed, increasing punitiveness appears to be, with isolated regional exceptions (within the United States, as well), a more general feature of modern Western society, exhibiting cross-national differences that are largely in degree rather than in kind (see Tonry & Farrington, 2005, and the chapters therein).

One social disability that has garnered sustained research attention is the employment barrier. Many ex-inmates report feeling that their criminal record hinders their ability to find a job (Visher, Debus-Sherrill, & Yahner, 2011), while experimental audits confirm that employers are only half as likely to call back job applicants who report a prison sentence on their application (Pager, 2003). Furthermore, in comparisons between ex-inmates and comparable non-incarcerated individuals, ex-inmates consistently exhibit employment probabilities that are about 10-15 percent lower (Apel & Sweeten, 2010; Huebner, 2005; Waldfogel, 1994). Among employed ex-inmates, there is a comparable earnings penalty on the order of 10-15 percent, as well as modestly slower earnings growth over time (Apel & Sweeten, 2010; Waldfogel, 1994; Western, 2002).

The importance of the foregoing findings lies in the widespread expectation that efforts to improve ex-inmates' success in the labor market can lower the risk of criminal behavior, and at the same time give them the capacity to earn a livable wage that further lessens the attraction of illegal behavior. Unfortunately, compared to the general population, ex-inmates possess deficits that would greatly limit their employment prospects, even in the absence of imprisonment. For example, they are overwhelmingly drawn from socially marginalized populations – the poor, minorities, high-school dropouts – and they tend to have erratic work histories (Dumont et al., 2012; Raphael, 2011; Wakefield & Uggen, 2010; Western, 2006). In spite of these obvious disadvantages, employment opportunities early in the prison reentry process do in fact have the capacity to strengthen commitments to conformity and to hasten desistance from criminal behavior (Uggen, 1999, 2000; Uggen & Thompson, 2003).

The current study seeks to fill in two important gaps in existing research, by analyzing differences in employment outcomes by imprisonment length in a sample of Dutch offenders who experience comparatively short prison

spells.[1] First, virtually no research attention has been devoted to sentences of incarceration less than one year. This is a glaring omission, as an example from the U.S. context makes clear. The jail population on any given day tends to be about one-half the size of the prison population – the prison incarceration rate is about 500 per 100,000, and the jail incarceration rate is about 250. However, the average daily jail population of about 750,000 individuals, from which the jail incarceration rate is calculated, underestimates by a very large margin the number of individuals who actually pass through the nation's jails in a year, which the Bureau of Justice Statistics estimates at a shade under 13 million in 2010 (Minton, 2011). Accounting for the fact that just under 40 percent of individuals are in jail serving a sentence (rather than awaiting trial) still yields about 5 million who are incarcerated in jail for a crime in a given year (Minton, 2011). So while the prison incarceration rate of about 500 per 100,000 provides a fairly accurate estimate of the number of people who spend time in prison during a year, the annualized jail incarceration rate actually exceeds 1,600 per 100,000 (or an alarming 4,300 per 100,000 if jail incarceration and pretrial detention are both considered). Short sentences of incarceration are clearly the norm, in the United States and elsewhere, yet virtually nothing is known about their consequences.

Second, an analysis of imprisonment length among incarcerated offenders allows one to isolate processes of erosion that are independent of the stigmatizing potential of incarceration. Estimates of the impact of incarceration (vs. non-incarceration) confound demand- and supply-side processes, which can cloud interpretation and impede the creation of policy solutions. Stigma is the quintessential demand-side impact of incarceration (see Holzer, Raphael, & Stoll, 2006). Yet limiting attention to incarcerated subjects only, the stigma of incarceration is held constant, at least in principle.[2] Any differences in employment prospects which are attributable to imprisonment length can then be interpreted as productivity losses that accrue as individuals spend more time isolated from the formal labor market.

In what follows, the prior literature on the incarceration-employment relationship is first reviewed. The discussion then turns to an overview of the Netherlands, the unique social context in which this study is carried out. Thereafter, the data and methods are described, followed by the results and an extended discussion.

1 In this study, imprisonment length is intended to refer to the actual length of time served, rather than to the sentence length handed down by a judge.

2 In fact, stigma could depend on the length of incarceration if longer imprisonment spells are more difficult to hide from employers.

4.2 Theoretical explanations

A significant policy concern is the degree to which incarceration stigmatizes ex-prisoners in the marketplace. The most direct evidence for stigma is provided by the hiring preferences and practices of employers. For example, Holzer et al. (2006) find that employer willingness to hire known ex-offenders (38% "probably will" or "definitely will" hire them) for low-skill positions is markedly lower than their willingness to hire welfare recipients (92%), applicants with only a GED (96%), applicants who had been unemployed for a year or more (83%), and applicants with a spotty work record (59%). Apparently, the employment prospects of ex-prisoners are negatively impacted by demand-side preferences which penalize individuals with a criminal history.

More relevant for the present study are the potential supply-side deficiencies that accumulate among offenders who are confined for longer periods of time – specifically, productivity losses due to longer imprisonment length. Three such processes include erosion in work skills, deepening embeddedness in illegal activity, and growing detachment from the institution of work – these are likely to be reinforcing rather than mutually exclusive processes. First, the most obvious source of erosion is depreciation in "human capital" as work-related skills and experiences go unused. The work histories of ex-prisoners will be punctuated with unaccounted-for gaps. Recall from Holzer, Raphael and Stoll's work (2006) that employers are only modestly more enthusiastic about hiring applicants with a spotty work record as they are about applicants with a criminal record (59% vs. 38% would hire them, respectively), compared to other difficult-to-employ groups (each in excess of 80% hiring likelihood). Therefore, even an employer who is completely unaware of an applicant's imprisonment will be reluctant to hire him over other unskilled individuals simply because the lack of stable work experience might convey that the applicant is unskilled, unreliable, or difficult to work with.

Second, longer confinement could also promote the accumulation of "criminal capital," or criminal knowledge and experiences which improve one's prospects in the illegal market (Hagan, 1993). Offenders can become more deeply embedded in criminal contexts as they are isolated from conventional society for longer periods of time, perhaps because they spend more time in the company of fellow captives who might strengthen their orientation to unlawful behavior. For example, Bayer, Hjalmarsson, and Pozen (2009) report reinforcing effects on recidivism exhibited by juveniles who are exposed to other youthful offenders remanded to the same correctional facility.

Third, longer imprisonment might also weaken an offender's attachment to legal work. Apel and Sweeten (2010) demonstrate that following their first conviction, young people who are incarcerated spend significantly more time "out of the labor force," relative to comparable young people who are not incarcerated following their first conviction. In other words,

incarceration is followed by a period of time in which ex-inmates are neither employed nor seeking employment (in contrast to unemployment, defined as non-employment but active job search). This work detachment endures for up to six years following confinement. While work detachment could partly reflect discouragement (labor force dropout precipitated by failed job search), the work histories of the to-be-incarcerated youth in this study were already characterized by longer periods of labor force non-participation. The confinement experience apparently worsened an already tenuous attachment to legal work.

To be sure, there are also reasons to believe that longer imprisonment length could actually increase employment prospects. Namely, longer periods of incarceration can increase a prisoner's exposure to correctional rehabilitation programs focusing on educational certification and vocational training, or to prison labor programs that provide tangible work experience. And there is suggestive evidence that such programs are effective in improving employment and lowering recidivism (Wilson, Gallagher, & MacKenzie, 2000). Although the commitment to rehabilitation has been more "rhetoric than reality" in U.S. prisons (Phelps, 2011), the same is not true in other Western contexts. Rehabilitation was a major punishment goal in the Netherlands after World War II. In the decades that followed, this focus became increasingly subordinate to other tasks of the prison system (e.g., humane detention and cost-effectiveness) (Downes & Van Swaaningen, 2007). Still, to the extent that the correctional system in the Netherlands has adopted a culture of rehabilitation, we regard any estimates of the relationship between imprisonment length and employment among Dutch ex-inmates as highly conservative.

4.3 Previous research

Limited empirical evidence actually exists on the question of the impact of imprisonment length on employment prospects, although two kinds of studies can be identified. The first set of studies represent analyses in which imprisonment length is not necessarily the primary determinant of employment under consideration. This includes a reanalysis of the Transitional Aid Research Project (Needels, 1996), a reanalysis of the National Supported Work Demonstration (Matsueda, Gartner, Piliavin, & Polakowski, 1992), and a reanalysis of data from a sample of males sentenced to a Boston-area reform school who were matched to school-going youth (Sampson & Laub, 1993). In these studies, the findings about the salience of imprisonment length are frustratingly mixed. Notably, however, they focus not on imprisonment length for any particular spell, but instead on total time incarcerated within a reference window, which can (and often does) include more than one incarceration spell.

A second prominent strand of scholarship that is directly focused on imprisonment length uses data from state correctional databases and unemployment insurance (UI) systems for contemporary samples of prisoners. Studies using administrative data have been conducted in Florida (Kling, 2004, 2006), Ohio (Sabol, 2007), Washington State (Pettit & Lyons, 2007, 2009), and Illinois (LaLonde & Cho, 2008; Jung, 2011). These studies consistently report that employment and earnings in UI-covered jobs, following release from prison, increase in sentence length. Kling (2004, 2006) finds that employment rates among ex-prisoners in Florida peak immediately upon release, at 40 percent among ex-prisoners incarcerated for one year, but over 50 percent among ex-prisoners incarcerated for four years (2006: 867, Figure 1A). Similarly, mean earnings in the peak quarter are about $800 among ex-prisoners incarcerated for one year, but $1,600 among ex-prisoners incarcerated for four years (Figure 1C). On the other hand, the differentials appear to be relatively short lived, as employment and earnings converge after about two years have elapsed. Furthermore, post-prison employment rates eventually return to their pre-prison level, irrespective of imprisonment length.

Pettit and Lyons (2007, 2009) similarly report that imprisonment length is positively and significantly correlated with employment rates among ex-prisoners in Washington State. As in the Florida study, there is a tendency for employment rates to return to pre-prison levels within about two years. Jung (2011) reports the same kind of convergence at the two-year mark among male ex-prisoners in Cook County, Illinois (see LaLonde & Cho, 2008, for evidence on female ex-prisoners in the same jurisdiction). Also like the Florida study, there are significant earnings differentials in favor of employed ex-prisoners who serve longer terms of confinement in Washington State and Illinois. Unlike the Florida study, on the other hand, while the earnings differentials narrow over time, they nevertheless appear to persist for the duration of the follow-up period. Yet the long-term differentials are not particularly notable for their magnitude. For example, in Jung's (2011) study, ex-prisoners who differ by one year in their imprisonment length only differ by about $150 in long-run quarterly earnings, or $50 per month (2011: 513, Table 4A, Model 6).

To summarize the administrative studies of imprisonment length and employment, the findings are unambiguous that offenders who serve longer prison terms have better prospects with respect to both employment and earnings (conditional on employment). However, these differentials tend to erode with the passage of time. A notable feature of these studies is that the subjects are state prison inmates, almost all of whom serve sentences longer than one year. In fact, the average ex-prisoner in these studies serves about two years behind bars. The estimates should therefore be interpreted as the correlation between imprisonment length and employment prospects, conditional on serving a prison sentence of well over one year.

4.4 Limitations of previous research

Two comments about this literature are in order, which help motivate the current study. First, administrative earnings data come from state tax records and are based on the earnings reported by employers to the state unemployment insurance (UI) system, and therefore fail to capture income from uncovered jobs (i.e., self-employment income, out-of-state income), among other sources of error. Comparisons of self-report and administrative data show that survey earnings are routinely higher than UI earnings, although program impacts tend to be similar (Kornfeld & Bloom, 1999). The one noteworthy exception is for young males with a criminal record, for whom the discrepancy between unofficial and official earnings is greatest, and for whom program impact estimates qualitatively differ depending on the source (Kornfeld & Bloom, 1999; Schochet, Burghardt, & McConnell, 2008). These are precisely the subjects of interest in the studies cited above, suggesting that post-prison employment prospects measured from tax records miss many sources of income for high-risk samples – self-employment, informal employment arrangements, short-term employment, and employment that is cash only or "off the books." If the tendency to work in UI-covered jobs varies systematically by the length of time served in prison, the positive correlation between imprisonment length and employment prospects reported in the studies cited above will partially be an artifact of this tendency. This suggests that self-report employment and earnings, while undoubtedly subject to their own peculiar sources of measurement error, are likely to be less biased for young, high-risk males and more generally, people with criminal records.

Second, most previous research does not consider the process by which incarceration shapes the work prospects of ex-prisoners. Administrative datasets, in particular, are not well suited to an elaboration of potential erosion processes that are linked to imprisonment length. In this study, we consider a number of measures to characterize the post-prison work experience: the timing of job acquisition, multiple job holding, wages, occupational level, and re-employment with a pre-prison employer. We also consider non-employment measures such as skills acquisition in prison and criminal recidivism. The analysis to follow is therefore capable of considering job stability, job quality, human capital, and criminal embeddedness.

4.5 The Netherlands as context

The United States is quite unique in its scale of imprisonment, but penal punitiveness is a much broader Western phenomenon (Tonry & Farrington, 2005). Two-thirds of 35 European countries surveyed recently have experienced prison growth (Aebi et al., 2006). The Netherlands in particular, long known for its liberal penal policies, has witnessed rapid prison expansion, growing almost fourfold (375 percent) during the last three decades (see

Tonry & Bijleveld, 2007). As a point of comparison, from 1975 to 2005, the U.S. incarceration rate grew 5.5 percent annually, while the comparable figure for the Netherlands is 4.9 percent.

Despite comparable prison growth rates, there are obvious differences between the Dutch and American penal climates. First, the Netherlands has a milder penal climate which might make the transition from prison to the labor market less fraught. Over 80 percent of all prisoners released in the Netherlands are confined for a maximum of six months. The median time served is one month and an average prison spell is 3.6 months (109 days) (Linckens & De Looff, 2012). State prisoners in the United States serve an average sentence of two years (Guerino et al., 2011). Moreover, prison conditions are generally less harsh in the Netherlands. For instance, most Dutch (pretrial) prisoners are confined in single cells in comparatively small prisons and entitled to daily yard time.

Second, criminal records are not publicly accessible in the Netherlands, and employers have few avenues to retrieve this information. Yet, in certain sectors, a conduct certificate is mandatory (e.g. education, health services, cab driving, security and logistics) and the rules for granting a certificate have become stricter, although it contains no information regarding the existence or nature of an applicant's criminal history (see Boone, 2011). These regulations aim to protect Dutch ex-offenders from labor market discrimination, whereas open access laws in many states in the United States pose an additional burden for American ex-offenders (see Briggs et al., 2004; Bushway, 2004).

Third, the more generous social welfare regime in the Netherlands might actually serve as a disincentive for employment among Dutch ex-offenders. Individual responsibility is not as strongly stressed (Becker, 2000; Esping-Andersen, 1990), and despite retrenchment in recent decades, the Dutch welfare system is still very generous compared to that of the United States. For example, in 2009 the Dutch government spent 23 percent of its national GDP on public expenditures (e.g. unemployment, housing, labor market programs, pensions), compared to 19 percent in the United States (OECD Social Expenditure Database (SOCX)). While this difference is substantial in itself, higher income inequality (in 2009 the Gini coefficient was 0.29 in the Netherlands and 0.38 in the United States) and demand for social provisions in the United States further emphasize the differences in social policy. On the other hand, higher minimum wages in the Netherlands might also lead to relatively higher employment rates among Dutch ex-prisoners. In addition, social benefits might provide some basic needs which make it easier for Dutch ex-prisoners to find and hold down a steady job.

Given these seemingly fundamental differences, one might reasonably ask whether findings from the Netherlands provide any generalizability at all to the American context. Where basic criminological relationships are concerned, this can be answered in the affirmative. The findings from many prior Dutch studies confirm the relevance of life transitions such as educational attainment, employment, marriage, and parenthood for crimi-

nal behavior and desistance (Blokland & Nieuwbeerta, 2005; Ramakers, Bijleveld, & Ruiter, 2010; Van der Geest, Bijleveld, & Blokland, 2011; Van Schellen, Apel, & Nieuwbeerta, 2012), as well as the impact of incarceration on marital stability and subsequent offending (Apel, Blokland, Nieuwbeerta, & Van Schellen, 2010; Nieuwbeerta, Apel, & Blokland, 2009; Snodgrass, Blokland, Haviland, Nieuwbeerta, & Nagin, 2011). However, because Dutch prison sentences are so much shorter than their American counterparts, it is best to conceive of them as being more akin to short incarceration spells, most of which are spent in jail. An analysis of imprisonment length and employment prospects in the Netherlands can thus fill an important empirical gap where short sentences of incarceration are concerned.

4.6 Data

The data for this study were collected as part of the Prison Project, a unique, longitudinal and nationwide effort to collect data about Dutch prisoners in the beginning of pretrial detention, during confinement, and after release from prison. The project targeted male prisoners who entered a Dutch detention facility between October 2010 and March 2011, were born in the Netherlands, were between 18 and 65 years old, and did not suffer from severe psychological problems.[3] A total of 2,945 prisoners who entered pretrial detention between October 2010 and March 2011 met the selection criteria. No less than 95 percent of these men were approached and 65 percent of them agreed to participate in the study. This sample of 1,909 prisoners was generally representative of all prisoners that met the selection criteria in terms of age, marital status, committing a violent crime, and receiving an unconditional prison sentence for the criterion offense, but differed slightly in some other characteristics.[4]

The sample used in the current study comprises 702 ex-prisoners who participated in the in-prison interview (P1) and agreed to a reentry interview (R1) that was conducted six months after release (up to June 2012).[5]

---

3 The study targeted prisoners who were detained for a minimum of 3 weeks. In some cases, prisoners were interviewed really soon after entering pretrial detention, but released shortly after this interview

4 Participants and non-participants differ with respect to age of onset (18.9 vs. 17.4), being employed before imprisonment (45.7% vs. 38.7%) and duration of actual time served (5.1 vs. 4.1 months). In addition, a comparison of criminal histories revealed that participants have a slightly less extensive criminal history than non-participants (7.7 vs. 9.8 previous convictions; 3.4 vs. 5.0 previous prison spells).

5 The R1-interview took place in prison if the subject had re-entered prison at that time (approximately 20% of the R1-interviews were held in prison). Not all interviews took place exactly six months after release from prison. A majority were held 5 or 6 months after release (63.6%), 2.6 percent of the interviews took place before that time and 33.8 percent took place at a later time.

As expected, the particular lifestyle of the sample made it difficult to contact the 1,423 ex-prisoners who were eligible for participation in the R1-interview. Some refused permission to be approached in follow-up waves (n = 43). Still, 76 percent of the 1,380 ex-prisoners were successfully contacted and 52 percent of them agreed to participate in the R1-interview. This led to an overall response rate of almost 34 percent (P1:0.65 x R1:0.52) in June 2012.

Importantly, difference tests showed comparability between the R1- and P1-samples across a wide range of baseline covariates (e.g, criminal history, parenthood, employed before imprisonment, educational level). Nevertheless, the R1-sample contained fewer non-ethnic Dutch (35% vs. 49%), fewer respondents with partners (44% vs. 53%) and fewer violent offenders (42% vs. 47%) than the full P1-sample. In addition, since not all 1,909 prisoners were eligible for participation in R1 (some were still imprisoned or had not yet been released for six months), the R1-sample at present has served a shorter prison spell (median spell of 3.1 months vs. 3.8 months). Some caution is therefore advised when generalizing the results from the R1-sample to the larger population of detainees from which the P1-respondents were drawn.

4.7 Measures

4.7.1 Length of imprisonment

Length of imprisonment is measured as the actual time between the first day of pretrial detention and the date of release from confinement (either pretrial detention or imprisonment), as registered by the Judicial Institutions Department of the Netherlands (mean = 3.8 months; median = 3.1 months, see Figure 4.1). In order to estimate the impact of different "doses" of imprisonment on employment, five groups are created: 1 to 6 weeks (n = 132), 6 weeks to 3 months (n = 191), 3 to 4 months (n = 133), 4 to 6 months (n = 127) and 6 to 12 months (n = 119). These are based on the distribution of length of imprisonment as well as judicial practice (i.e. how decisions about extended placement are made by judges) (Tak, 2003). In follow-up analyses detailed in the appendix, imprisonment length is retained as the number of days of confinement (continuous measurement).

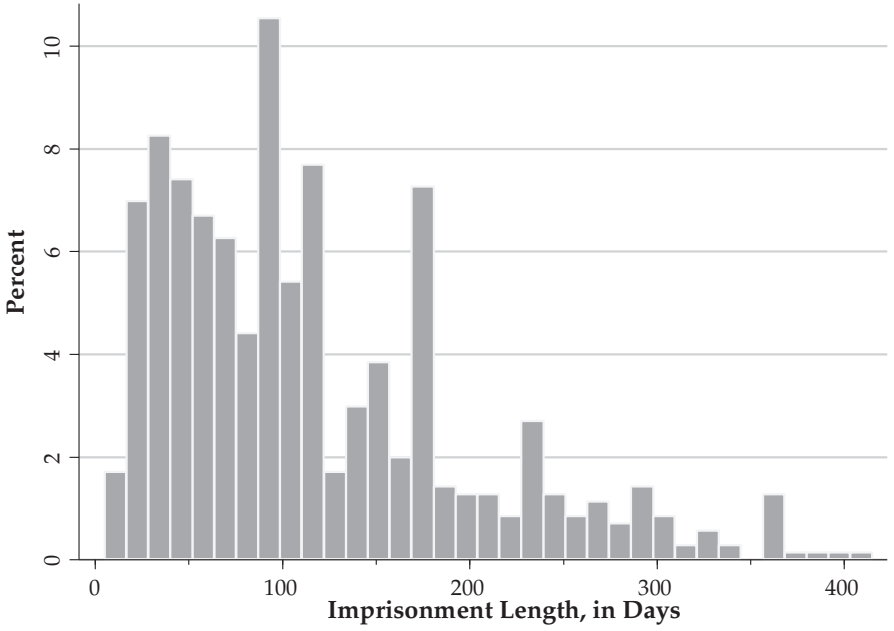


Figure 4.1 Distribution of imprisonment length (n = 702)

4.7.2 Post-prison outcomes

Employment during the six months post-release is measured in two distinct ways during the R1-interview. The first measure is from survey responses to the question, "How many jobs did you have since leaving prison?" The second measure is from a life event calendar, in which respondents report on their employment situation (among other information) during each month since their release from prison. Previous research has shown that this method leads to higher quality retrospective reports (Belli, Shay, & Stafford, 2001; Engel, Keifer, & Zahm, 2001). Respondents who report being self-employed, being an employee, or working "off the books" in a given month are counted as "employed" (months spent in prison were counted as unemployed months). Both the survey and calendar measures indicate that roughly one-half of respondents find employment within six months of leaving prison (see Table 4.1).

In addition to these binary indicators of employment, we construct four measures of job stability. We identify ex-prisoners who work in more than one job, and we use the life event calendar to compute the proportion of months employed during the six-month reentry window. Additionally, we determine whether employment was found immediately upon release, and identify the employed ex-prisoners who returned to their pre-prison job after release.

We also utilize two measures of job quality. At the R1-interview, employed respondents report their net monthly salary (after taxes) (mean = €2,109; median = €1,500).[6] In addition, based on the Standard for Classification of Occupations (SBC) of Statistics Netherlands (Westerman, 2010), information about the job title, type of business, and (executive) tasks was used to classify self-employed and salaried workers into one of five occupational levels: elementary, low, middle, high, or scientific.

Finally, we include two non-employment outcomes – human capital and criminal capital – to gain insight into the processes of skills erosion and criminal embeddedness that might underlie the effect, if any, of imprisonment length on employment. Human capital accumulation is measured by participation in a prison program. Criminal embeddedness is based on self-reports from the life-event calendar about whether respondents have committed a crime since their release from prison.[7]

Table 4.1 Descriptive statistics

	N	Mean	Median	SD	Min	Max
Found employment within six months [a]	694	.51			.00	1.00
Found employment within six months [b]	651	.47			.00	1.00
Worked multiple jobs within six months [a]	351	.24			.00	1.00
Employed in 1st month [b,c]	308	.67			.00	1.00
Time spent employed [b,c]	308	79.32	100.00	27.03	16.67	100.00
Return to pre-prison employer [d]	188	.30			.00	1.00
Wage in sixth month (€) [e]	236	2,109.30	1,500.00	4,222.09	10.00	58,000.00
Occupational level in sixth month [e]	236	2.20	2.00	.55	1.00	5.00

ABBREVIATIONS: SD = standard deviation (omitted for dummy variables), Min = Minimum, Max = Maximum.

[a] Based on general question concerning the number of jobs during the follow-up period.

[b] Based on monthly employment data from the calendar questionnaire.

[c] Available for those employed (for at least one month) during the follow-up period.

[d] Available for those who were employed as salary workers before imprisonment.

[e] Available for those who were employed in the sixth month after release.

4.7.3 *Pre-prison confounding variables*

The background data collected in the Prison Project are quite rich, allowing us to eliminate a wide range of potential pre-existing differences between men who serve different lengths of imprisonment. Appendix 4.A contains descriptive information on 55 such covariates. This information is incorpo-

6 All values above the 95-percentile were truncated.

7 Additional analyses (not shown) indicated that a difference in exposure time (time spent in prison during the six-month follow-up) could not explain the lower employment ratio among long-term prisoners. In addition, we ruled out the possibility that long-term prisoners were less likely to find employment because they were more likely to get sentenced back to prison for the criterion offense – among those who were released before trial, only six long-term prisoners returned to prison for the criterion offense.

rated into a propensity score model, and includes demographics, employment history since leaving full-time education, employment situation before imprisonment, social bonds, sources of income before imprisonment, lifestyle, and attitudes. The Public Prosecutor's Office was consulted for information on the "criterion" offense: the number of registered offenses in a criminal case, the maximum penalty (maximum days a judge can sentence an offender to prison based on the criterion offense), and pretrial release. Detailed information on the type of crime and the offender's criminal history was collected from "rap sheets" available in the Criminal Record Office. These data were made available by the Research and Documentation Centre (WODC) of the Netherlands Ministry of Security and Justice, and contain information on all registered convictions beginning at age 12, the age of criminal responsibility.

4.8 Analytical approach

The main objective of this study is to identify the effect of imprisonment length on employment outcomes among pretrial detainees who were detained for a minimum of one week. A simple comparison of post-prison employment rates across groups that served different lengths of imprisonment is potentially confounded with pre-prison factors that affect not only the length of imprisonment but also labor market performance. Only an experimental design, in which individuals are randomly assigned to prison for shorter or longer periods of time, would ensure that all possible confounders (including unobservables) are controlled. However, any bias caused by observable pre-prison covariates can be eliminated by conditioning on a propensity score (Rosenbaum, 2002; Rosenbaum & Rubin, 1983). A general advantage of the propensity score methodology over standard regression analyses is that it is more robust with respect to model misspecification (Drake, 1993). Another advantage is the internal validity that results from this approach, as it assures the exclusion (or down-weighting) of "treated" individuals for whom no comparable "controls" are available.

The richness of the Prison Project data and large sample size allow us to rule out 55 potential confounders. To our advantage, most covariates (47 out of 55) are initially balanced, which indicates that groups which differ in imprisonment length are already highly similar (see Appendix 4.B). Failure to account for the remaining observable differences would allow selection bias to contaminate the results. And, to be sure, there may still be hidden biases confounding our results after the differences in observables are taken into account (Rosenbaum, 2002). Yet, a substantial share of the potential confounders can be eliminated using propensity score methods. We account for factors that have been shown to be highly influential in sentencing decisions (crime severity, criminal history, demographics), and the rich background data enable us to incorporate many more.

4.8.1 *Propensity score model for an ordered treatment*

A propensity score is a type of "balancing score" which represents the probability of receiving treatment, conditional on a set of observed pre-treatment covariates. In the case of a binary treatment, two individuals with an identical or closely similar propensity score, but a different observed treatment, are compared in outcome. In this study, however, subjects served different imprisonment lengths and were classified into one of five groups receiving smaller or larger "doses" of prison (1 = 1-6 weeks, 2 = 6 weeks to 3 months, 3 = 3-4 months, 4 = 4-6 months, 5 = 6-12 months). Following Loughran et al. (2009), the current study uses sub-classification on the balancing score. This score was estimated from an ordered logit model in order to create groups of prisoners who are observationally similar on measured covariates at the time of arrest, yet served different lengths of confinement (for other applications of the generalized propensity score, see Lu et al., 2001; Zanutto, Lu, & Hornik, 2005). The probability that subject i serves imprisonment length D_i or higher, conditional on $j = 1,\ldots,K$ pre-treatment covariates X_{ij}, is parameterized in familiar log-odds form:

$$ln\left[\frac{Pr(D_i > d)}{Pr(D_i \leq d)}\right] = \tau_d + \sum_{j=1}^{K} \beta_j X_{ij}, \text{ for } d = 1,2,3,4$$

In this model, τ_d represents a dose-specific threshold or intercept, corresponding to imprisonment length d (exclusion of the threshold for one category, $d = 5$ in this case, is necessary for identification). By satisfying the proportional odds assumption, a single set of coefficients can be estimated for each of the covariates, ensuring that the only difference in the likelihood of different imprisonment lengths is an intercept shift captured by the thresholds.[8] The implication is that the thresholds, because they are constants, can be removed and a single balancing score estimated for each subject using just the linear predictor from the ordered logit model.

This balancing score is the ordered logit analog to the propensity score estimated from a binary logit model. One difference is that, because the balancing score in this analysis is used for stratification rather than matching, no additional transformations are necessary (e.g., conversion of the balancing score to a probability). The balancing score is used to group subjects into five equal-sized subgroups, known as strata, within each imprisonment length group. Classification into five strata suffices to remove approximately 90 percent of the initial imbalance in each of the covariates (Rosenbaum & Rubin, 1984). Covariate balance is maximized using an iterative approach for model selection (e.g., including interactions, squares, log transformations), after which the stratum-weighted mean of employment outcomes, conditional on receiving imprisonment length d, is estimated (see also Loughran et al., 2009).

8 The proportional odds assumption is easily satisfied (χ^2 [180] = 188.9, p = 0.31).

4.8.2 Propensity score estimation and covariate balance

Our objective is to compare the post-prison employment outcomes of individuals who are observationally similar with respect to pre-treatment covariates (as indexed by the balancing score), yet served different lengths of imprisonment. We take advantage of the overlap in predicted balancing scores across the five groups of detainees (see Figure 4.2) and exclude the cases for which no appropriate match is available. Hence, the analytic sample excludes subjects who have a balancing score lower than the minimum score among the long-term prisoners (n = 7), or a balancing score higher than the maximum score among the short-term prisoners (n = 29) (see the black dashed lines in Figure 4.2). The final analytic sample thereby consists of 666 subjects.

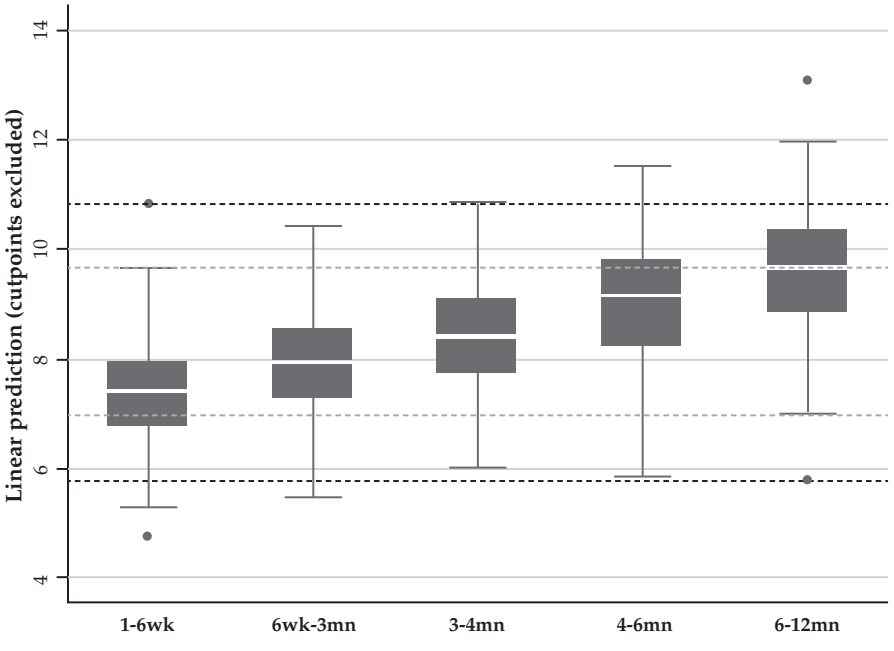


Figure 4.2 Balancing score distributions, by imprisonment length (n = 702)

Note: The black dashed lines indicate the trimming performed prior to stratification for the main analysis, specifically, propensity scores lower than the maximum of the shortest imprisonment length group and higher than the minimum of the longest imprisonment length group were included. The gray dashed lines indicate the trimming performed as a sensitivity analysis, specifically, propensity scores lower than the upper adjacent vazlue (i.e., the upper whisker) of the shortest imprisonment length group and higher than the lower adjacent value (i.e., the lower whisker) of the longest imprisonment length group.

Following Loughran et al. (2009), covariate balance was evaluated by performing two-way analyses of variance (ANOVA). The ordinal measure of imprisonment length, the balancing score strata, and their interaction serve as independent variables, where each covariate is the dependent variable.

Balance is assured when the combination of imprisonment length and its interaction with the balancing score strata is not significantly correlated with the covariate ($\alpha = 0.05$). After stratification, age of onset, sex crime as criterion offense, and ethnic background remain out of balance (see Appendix 4.C), but we would expect about 3 of the 55 covariates to be out of balance by chance alone.[9]

More recently, scholars have discouraged the use of significance tests to check balance, because these tests can be affected by not only changes in effect size but also changes in sample size (see Connelly, Sackett, & Waters, 2013). Therefore, the magnitude of group differences in covariates was assessed by performing regressions with the covariate as the outcome and the ordered treatment as the regressor, weighted by the propensity score strata. The square root of the R-square from this model functions as a measure of effect size (0.1 = small; 0.3 = medium; 0.5 = large). A covariate is considered to be out of balance when this effect size is 0.10 or higher. These tests show that the relatively large group differences in several of the criterion offense characteristics (between 0.16 and 0.36) decreased after stratification (highest effect size is 0.12) (see Appendices 4.B and 4.C). Of the three covariates that were out of balance based on the aforementioned significance tests (age of onset, sex crime as criterion offense, ethnic background), two had very small effect sizes before and after stratification.

We also find that some effect sizes actually increased after stratification (e.g., the effect size for use of alcohol and sex crime increased from 0.02 to 0.12). Nevertheless, Appendix 4.C shows that most effect sizes decreased after stratification and are far below the threshold of 0.10,. Hence, even though the propensity score method used here does not enforce complete balance between imprisonment length groups, both the significance tests as well as the effect size measures indicate that the current model confronts the selection problem as rigorously as possible by eliminating a substantial number of covariates as potential confounders.

4.9 Results

4.9.1 Finding employment

This section presents the adjusted findings (weighted by propensity score strata). The key finding with respect to the impact of imprisonment length on employment is shown in Figure 4.3. In each month after release, men who stayed in prison for more than six months have a lower likelihood of employment than their observationally similar counterparts who stayed in

9 Sensitivity analyses on the effect of imprisonment length on employment likelihood in which we directly adjusted for the influence of the out-of-balance covariates – by including them together with imprisonment length in a regression model weighted by the propensity score strata – led to similar conclusions.

prison for less than six months. For instance, the sample average employment likelihood in the first month after release is 31 percent, but the long-term prisoners (minimum spell of six months) exhibit a substantially lower employment likelihood (20%). Though not linear, the relationship between imprisonment length and first-month employment is marginally significant ($\chi^2 = 7.91$, $p < 0.10$; Cramer's V = 0.113). This difference in employment likelihood does not remain intact during all follow-up months, but standardized residuals indicate that the employment rate of long-term prisoners remains significantly lower throughout the first six months after leaving prison. Additional analyses of the type of employment (not shown) indicate that long-term prisoners are less likely to work in a formal job (as salary workers) and to be self-employed than short-term prisoners.

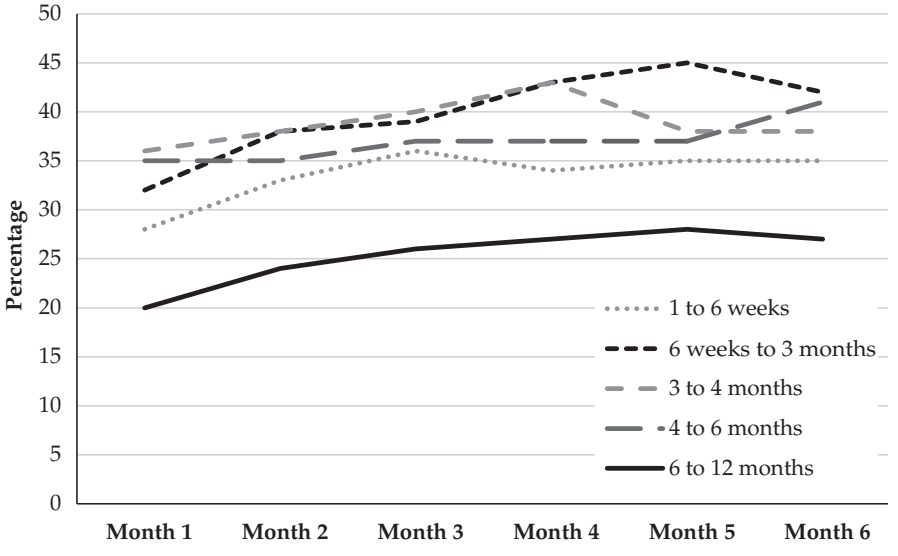


Figure 4.3 Monthly employment rates following prison release, by imprisonment length

Note: Figure 4.3 presents the adjusted stratum-weighted means (Panel B, Table 4.2). Employment is measured from the life event calendar.

Having examined the employment differentials in each month of the follow-up, we next compare the percentages of ex-prisoners who found employment at any point within the six-month window. Recall that we can use both the monthly calendar data as well as a general question about the number of jobs acquired since leaving prison. Table 4.2 provides the unadjusted estimates (Panel A) and adjusted estimates (Panel B). Both panels show a generally inverse relationship between imprisonment and employment. Even though the adjusted effect sizes are fairly modest for both measures (Cramer's V is 0.114 and 0.155), the significance of the monthly measure ($\chi^2 = 8.12$, $p < .10$) and general measure ($\chi^2 = 15.84$, $p < .01$) reveals that the diverging

impact of a longer prison spell on the overall employment rate remains intact after controlling for selection on observables. Since both measures are presumably subject to their own sources of measurement error, the high resemblance across measures strengthens the finding. They offer evidence for our expectation that ex-prisoners who stay in prison longer are less likely to be employed in the first crucial months following release, compared to their observationally similar counterparts who spend less time in prison. The strongest relationship is observed for the general question – 58 percent of the short-term prisoners (1-6 weeks) obtain a job, compared to 32 percent of the long-term prisoners (6-12 months), with the remaining groups intermediate between these two percentages at about 50 percent.

While the general employment measure exhibits a linear trend (χ^2 = 12.30, $p < .01$), the finding that stands out the most across both panels and measures is the relatively low employment likelihood of long-term prisoners (6-12 months). Additional difference tests between these long-term prisoners and all other prisoners confirm this, for both measures (see the bottom rows of Table 4.2). Specifically, the employment differential is –14 probability points (calendar measure) and –19 probability points (general measure) for offenders confined in excess of 6 months compared to all offenders confined for fewer than 6 months (Panel B).

4.9.2 Sensitivity analyses

In order to increase confidence in the findings thus far, three types of sensitivity analyses were performed. We performed the abovementioned analyses on a more restrictive sample, we estimated a propensity score model for a continuous treatment (number of days in prison) rather than ordered treatment, and we estimated a standard regression model controlling for the covariates directly rather than indirectly by way of a propensity score. Here we summarize the conclusions from these analyses.

First, we restricted the sample further to ensure even more similarity between offenders serving different confinement lengths. In the results reported above, we trimmed the sample to retain those with balancing scores lower than the maximum of the shortest imprisonment length group and higher than the minimum of the longest imprisonment length group (see the black dashed lines in Figure 4.2). In additional analyses, we instead trimmed the sample to retain those with balancing scores lower than the upper adjacent value (i.e., the upper whisker) of the shortest imprisonment length group and higher than the lower adjacent value (i.e., the lower whisker) of the longest imprisonment length group (see the gray dashed lines in Figure 4.2). This resulted in a substantially smaller estimation sample (N = 476 vs. N = 666).

Table 4.2 Post-prison employment

	Calendar questionnaire[a]			General question[b]		
	Unadjusted[c] (Panel A)	Adjusted[d] (Panel B)	(Panel C)	Unadjusted[c] (Panel A)	Adjusted[d] (Panel B)	(Panel C)
Imprisonment length	%	%	%	%	%	%
1 to 6 weeks	50.81	42.02	52.38	56.06	57.60	61.11
6 weeks to 3 months	48.31	46.07	49.64	50.80	49.19	49.30
3 to 4 months	50.00	49.58	47.87	51.91	50.39	49.51
4 to 6 months	46.34	42.86	46.58	51.59	47.01	53.16
6 to 12 months	39.62	31.18	42.86	42.37	31.68	42.11
All	47.31	43.16	48.51	50.72	47.80	51.38
Significance	NS	†	NS	NS	**	NS
Statistical test	$\chi^2[4]=3.59$	$\chi^2[4]=8.12$	$\chi^2[4]=1.33$	$\chi^2[4]=4.91$	$\chi^2[4]=15.84$	$\chi^2[4]=5.87$
Linear trend	No ($\chi^2[1]=2.59$)	No ($\chi^2[1]=2.29$)	No ($\chi^2[1]=1.28$)	Yes ($\chi^2[1]=3.31$)	Yes ($\chi^2[1]=12.30$)	Yes ($\chi^2[1]=2.91$)
N	651	621	437	694	657	471
6-12 months vs. all others						
Difference in mean (%)	-9.6	-14	-6.2	-10	-19	-10.6
Significance	†	*	NS	*	***	NS

NOTES: Panel A= whole sample (*n* =702), Panel B= main sample (after stratification on propensity score) (*n*=666), Panel C= alternative sample after stratification on propensity score (*n*=476).

ABBREVIATIONS: χ^2 = Chi square test.

[a] Based on monthly employment data from the calendar questionnaire.

[b] Based on general question concerning the number of jobs during the follow-up period.

[c] This column provides the unadjusted estimates (before stratification on propensity score)

[d] This column provides the adjusted estimates (stratum-weighted means)

†$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$ (two-tailed tests).

The results from the smaller, alternative sample are shown in Panel C of Table 4.2. Balance diagnostics indicated that after stratification, again, few covariates remained out of balance (namely, wage, self-employed before imprisonment, previous prison spell) and the magnitude of group differences decreased. All but one covariate exhibited an effect size below 0.10 (number of property crimes, $r = 0.10$). The main difference in findings is that Panel C indicates higher overall employment rates across all groups, implying that more of the "unemployable" ex-prisoners were excluded from this alternative sample. Nevertheless, the basic finding that a longer prison spell is correlated with lower employment chances is replicated for both employment measures.

Although the results in Panel C do not achieve statistical significance, the pattern is similar to that observed in Panel B – long-term prisoners (6-12 months) possess a much lower employment rate compared to short- and medium-term prisoners. For example, based on the calendar questionnaire, employment among offenders confined for 6-12 months remains about 6 probability points lower than offenders confined for less than 6 months. The employment differential is –11 probability points for the measure based on the general employment question. Interestingly, while these point estimates are halved from the Panel B results, they more closely resemble the unadjusted results in Panel A, which in the case of the general measure of employment, are statistically significant.

Second, we evaluated the sensitivity of the findings to the choice of propensity score method. In additional analyses, we considered a propensity score model for continuous treatment (number of days of prison confinement). Details on the approach are provided in Appendix 4.E, and here we briefly summarize the key findings. Figure 4.4 illustrates the mean probability of employment for specific imprisonment length "doses" (spanning 5 days to 415 days), conditional on the generalized propensity score (and thus the covariates indexed by it). Note that each subject contributes to the estimate of the mean probability evaluated at each imprisonment length "dose," as explained in Appendix 4.E. The graphs illustrate that the likelihood of employment is negative and mildly non-linear (but monotonic) in the length of prison confinement. For example, from 5 days to about 150 days (5 months), the mean employment probabilities decline with longer confinement, but tend to do so very slowly within this range; indeed, the confidence intervals overlap considerably. On the other hand, the dose-response function becomes more steeply and linearly inverse when imprisonment length exceeds about 180 days, or approximately 6 months. This harmonizes nicely with the results yielded by the ordered response model. It is also noteworthy that, as before, the findings from the general measure of employment indicate a stronger relationship than the findings from the calendar measure.

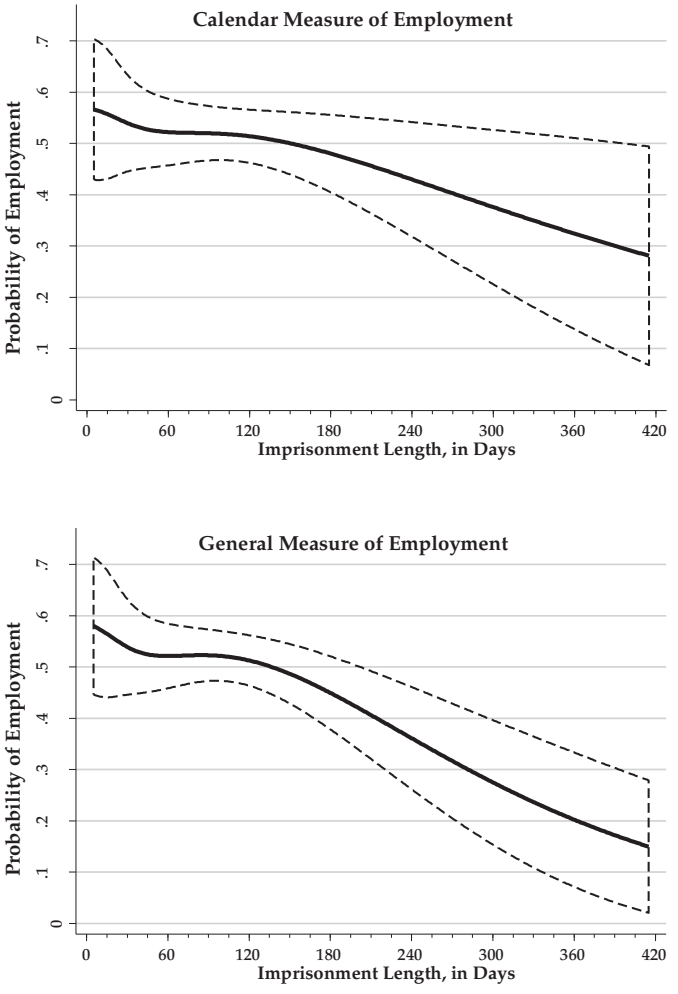


Figure 4.4 Dose-response relationship between imprisonment length (in days) and employment, from a generalized propensity score model

Note: The solid line represents the predicted mean probability of employment for a specific imprisonment length "dose," conditional on the generalized propensity score (and thus the covariates indexed by it). The dashed bands are bootstrapped 95% confidence intervals. Note that all 702 subjects contribute to the estimate of the mean for all imprisonment lengths between 5 days and 415 days, yielding 411 predicted probabilities per subject. Details on the generalized propensity score model and the creation of this graph are provided in Appendix 4.E.

Finally, in a series of models that are not shown, we controlled directly for the pre-prison covariates in a linear probability model of employment on imprisonment length (robust standard errors are used). For both employment outcomes, the coefficient for imprisonment length (in days) is –0.00075 ($p < .01$), indicating the impact of one additional day of confinement on

employment. Alternatively, when imprisonment length is logged, the coefficient from both models is about –0.080 ($p < .01$), which we can use to say that a doubling of imprisonment length (say, from 50 to 100 days, or from 100 to 200 days) corresponds with a decline in employment of 8 probability points, on average. Of course, these models impose a linear functional form to the relationship between imprisonment length and employment, whereas evidence reported above indicates a non-linear functional form. Squared terms in the models were not statistically significant, although a dummy variable for imprisonment length in excess of 6 months (relative to imprisonment less than 6 months) was marginally significant in the calendar employment model ($b = –.076$; $p < .10$) but non-significant in the general employment model ($b = –.050$; $p = .22$).

In summary, the main findings and sensitivity analyses yield evidence of a vaguely linear, inverse relationship between imprisonment length and employment, although the strength of the relationship is dependent on model choice and outcome measure. Although the evidence is not conclusive in all sensitivity analyses, the most consistent finding concerns imprisonment of 6 months or more: prison spells in excess of 6 months are correlated with diminished employment prospects after release.

4.9.3 Explaining the effect of imprisonment length on employment

In order to understand potential explanatory mechanisms for the employment differentials, we examine participation in prison programs and self-report recidivism. Table 4.3 shows that, even within a sample of relatively short prison spells, long-term prisoners have more opportunities to compensate for their absence from the labor market through educational programs and interventions in prison (χ^2=47.66, p <.01). This suggests that the employment differentials are likely to be even larger in the absence of in-prison programming. We also observe highly similar levels of recidivism across the groups: about one-fourth of the sample reports committing a criminal offense during the six-month follow-up. While not significantly different from the remaining groups, the long-term prisoners actually report the lowest recidivism rate at 17 percent. Thus, deepening embeddedness in criminal behavior seems incapable of explaining the lower employment rates among the ex-prisoners who serve the longest sentences.

Table 4.3 Human capital and criminal capital[a]

	Prison program participation	Recidivism [b]
Imprisonment length	%	%
1 to 6 weeks	3.20	21.01
6 weeks to 3 months	16.32	24.86
3 to 4 months	28.24	24.37
4 to 6 months	35.59	26.55
6 to 12 months	27.45	17.02
All	21.43	23.04
Significance	***	NS
N	666	620

[a] This table presents the adjusted stratum-weighted means of Panel B.
[b] Based on monthly self-report criminal behavior from the calendar questionnaire.
†$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$ (two-tailed tests).

4.9.4 Job stability and job quality

We examine the job stability of ex-prisoners who found employment in the first half-year upon release using four indicators: multiple job holding, immediate job acquisition, re-employment in a pre-prison job, and proportion of time spent employed. A quick glance at Table 4.4 shows that imprisonment length does not significantly impact any of the measures of job stability. A notable finding is that only 22 percent of employed ex-prisoners worked in more than one job, indicating low turnover. And as was already shown in Figure 4.3, the vast majority of employed ex-prisoners find a job immediately upon release (66-81%).

Table 4.4 Post-prison job stability[a]

	Worked multiple jobs [b]	Employed in 1st month [c]	Return to pre-prison employer [d]	Time spent employed [c]
Imprisonment length	%	%	%	Proportion
1 to 6 weeks	26.45	68.00	34.78	0.78
6 weeks to 3 months	22.84	69.14	26.67	0.85
3 to 4 months	19.44	74.14	28.21	0.77
4 to 6 months	19.45	81.25	31.03	0.84
6 to 12 months	17.70	65.52	20.00	0.80
All	21.85	71.80	27.92	0.81
Significance	NS	NS	NS	NS
Statistical test	$\chi^2[4]=1.39$	$\chi^2[4]=3.48$	$\chi^2[4]=1.51$	KW[4]=2.38
Linear trend	No ($\chi^2[1]=1.23$)	No ($\chi^2[1]=.545$)	No ($\chi^2[1]=.463$)	No (*J-T*=15477.0)
N	314	266	176	267

ABBREVIATIONS: χ^2 = Chi square test, KW = Kruskal-Wallis test, J-T= Jonckheere-Terpstra test (tests for an ordered pattern in the medians).
[a] This table presents the adjusted stratum-weighted means of Panel B.
[b] Based on general question concerning number of jobs and available for those who reported one or more jobs.
[c] Based on monthly employment data from calendar questionnaire and available for those employed for at least one month during the follow-up.
[d] Available for those who were employed as salary workers before imprisonment.

Approximately 28 percent of the previously employed prisoners were able to maintain their employment ties beyond their confinement, offering one explanation for quick job acquisition. Short-term prisoners were the most likely to return to a pre-prison employer (35%), and the long-term prisoners were the least likely (20%). While this difference is not significant, it does help partly explain the employment differentials by imprisonment length. Specifically, the relatively low employment ratio among long-term prisoners (6-12 months) seems to be driven by a combination of re-employment (short-term prisoners are more likely to return to their pre-prison job) and new job acquisition (short-term prisoners are more likely to find new employment). Further examination (not shown) points to self-employment as another plausible explanation for quick job acquisition: the majority of men who were self-employed upon release had also classified themselves as self-employed before imprisonment. Many of these men worked as independent contractors or owned small businesses.

The final indicator of job stability – proportion of time employed – suggests that those who find employment tend to remain employed for the greater part of the follow-up period, (0.77-0.85). Altogether these findings show that many of those who find employment are able to hold on to the same job, at least during the first half-year following release from prison.

Finally, we turn to the job quality of the men who were employed in the sixth month after release. Table 4.5 shows no significant differences in monthly earnings or occupational level between the imprisonment length groups. Compared to the average Dutch male worker who earns €2,275 per month (Statistics Netherlands, 2010), this sample averages earnings between €1,839 and €2,128.[10] For roughly one-third of the ex-prisoners, the income from employment is below the minimum monthly income in the Netherlands for adults (aged 23 and older), which is approximately €1,424 before taxes (assuming a 40-hours work week). Although not shown, the median hourly wage is approximately €9.37 and a third of the ex-prisoners earn below the legal minimum hourly wage for adults in the Netherlands (€8.22) (Ministry of Social Affairs and Employment, 2011). While these low earnings could be partly explained by the fact that 27 percent of the ex-prisoners have not reached adulthood (younger than 23), they nonetheless show that the ex-prisoners in this sample are concentrated in low-wage jobs.

10 As of this writing, €1.00 is roughly $1.35. Monthly earnings for this sample thus average between $2,483 and $2,873.

Table 4.5 Post-prison job quality[a]

	Wage (€) [b]		Medium / higher occupational level [b]
Imprisonment length	Mean	Median	Percent
1 to 6 weeks	1,907.44	1,500	38.46
6 weeks to 3 months	1,866.95	1,500	28.79
3 to 4 months	2,127.83	1,428	28.95
4 to 6 months	1,839.17	1,500	33.33
6 to 12 months	2,078.16	1,150	28.57
All	1,946.61	1,500	31.40
Significance	NS		NS
Statistical test	KW[4]=6.12		χ^2[4]=1.384
Linear trend	No (J-T=8376.50)		No (χ^2[1]=.282)
N	206		207

ABBREVIATIONS: χ^2 = Chi square test, KW = Kruskal-Wallis test, J-T= Jonckheere-Terpstra test (tests for an ordered pattern in the medians).

[a] This table presents the adjusted stratum-weighted means of Panel B.

[b] Available for those employed in sixth month after release.

We also consider the occupational level of post-prison employment by categorizing all jobs into "high status" (middle, high, or scientific level) versus "low status" (elementary or low level) occupations. Table 4.5 shows that less than one-third of the employed ex-prisoners obtain a high-status job. These men run their own (small) business (e.g., furniture, tanning studio, cars), for instance, or work as a manager, real estate agent, or landscaper. This contrasts sharply with 70 percent of the Dutch male work force that is employed in a high-status occupation (Statistics Netherlands, 2011). The individuals in low-status occupations do not direct other employees and work in jobs that require less education (e.g., warehouse worker, bicycle repairer, road worker). Hence, while many ex-prisoners seem able to find and hold down employment in the first crucial months after release, the quality of their post-prison employment tends to be quite low.

4.10 Discussion

4.10.1 Post-prison employment

About half the ex-prisoners found employment in the first six months after release, which studies indicate is a crucial window of time during reentry. The most salient finding was that ex-prisoners who were confined for longer than six months were less likely to be employed, compared to their observationally similar counterparts who were confined for a shorter length of time. When comparing these results with previous studies, we see resemblance with respect to employment rates. Administrative studies report post-prison employment rates of roughly 50 percent immediately after release (Sabol, 2007; Pettit & Lyons, 2007). Visher et al. (2011) found that 65 percent of ex-

prisoners were employed within eight months of release. On the other hand, our finding of a vaguely inverse relationship between imprisonment length and employment rates is not in line with other recent studies. For example, administrative studies consistently conclude that long-term prisoners are more likely to find employment and have higher wages in the first months following release than short-term prisoners (Jung, 2011; Kling, 2004, 2006; Pettit & Lyons, 2007, 2009).

One possible explanation for this contrast in findings is that many previous studies, because they are based on the use of administrative data, fail to capture income not reported to state unemployment insurance systems (self-employment, off-the-books employment, out-of-state employment), and this measurement error is likely to be most severe in samples of high-risk males (Kornfeld & Bloom, 1999; Schochet et al., 2008). Our survey data include all kinds of employment reported by ex-prisoners. Unfortunately, further analyses of the distinction between "formal" and "informal" employment in our data (not shown) were inconclusive, as few ex-prisoners reported working informally. This aligns with previous research showing that ex-prisoners might often fail to make a distinction between formal and informal work because they spend their whole working lives in the informal labor market (Fletcher, 2008).

Another explanation for the contrast in findings could be that our sample includes relatively short prison spells – much shorter than prior administrative studies, in which mean imprisonment length is two years. The negative impact of imprisonment on employment rates observed in our study is thus conditional on serving a prison sentence of well under one year. The apparent positive impact of longer prison spells observed in previous studies could be explained by stronger deterrent effects or more extensive institution- or community-based programming. Relatedly, the relative rarity of long prison spells in the Netherlands, or its more generous welfare regime, might also contribute to cross-national differences in patterns. Comparative research is warranted in order to test the validity of such explanations.

In analyses of non-employment outcomes, we found no evidence for the human capital or criminal embeddedness hypotheses – that skills erosion or recidivism among long-term prisoners explained the relatively low employment ratio among this group. In contrast, our results suggested that the differences in employment are likely to be even larger in the absence of prison programming. Deepening embeddedness in criminal behavior is also incapable of explaining the lower employment rates among long-term prisoners, as we found similar levels of criminal behavior across all imprisonment length groups.

4.10.2 Post-prison job stability and job quality

While we observed a lower employment ratio among the long-term prisoners (6-12 months), prison spells in excess of six months did not lead to different outcomes with respect to the measures of job stability (i.e., employment

timing, number of jobs, re-employment in a pre-prison job, and time employed). We recognize that our six-month window is too short to draw inferences about the long-term job stability of ex-prisoners, necessitating continued follow-up to ascertain whether ex-prisoners are able to keep their jobs for a long period of time.

It is noteworthy that the vast majority of employed ex-prisoners found a job quickly, and that pre-prison employment ties are of major importance for post-prison employment prospects. From a policy perspective, it might be fruitful to consider creative ways to incentivize employers to hire back former employees whose work was disrupted by a short prison spell, assuming of course that the criminal behavior which precipitated their incarceration was unrelated to work activities. Indirectly, this finding also reinforces the expectation that employment is far more difficult for those with little prior work experience, for whom *entry* rather than *reentry* into the labor market accurately characterizes the post-prison challenge (a point made by Bushway, 2006). With respect to job quality, the results show that the employed ex-prisoners often return to, or begin working in, uniformly low-quality jobs which differ little by imprisonment length. Our findings thus lead us to conclude that many of the employed ex-prisoners in our sample find jobs that are relatively stable but of uniformly low quality.

4.10.3 Limitations and future research

Some limitations of this study deserve attention in future research. First, it should go without saying that results from a propensity score model are only as strong as the covariates which are included in the analysis, and the method requires great care in the modeling of the "assignment mechanism." The propensity score methodology can only account for selection on observables, or measured differences between detainees prior to their incarceration. Our view is that the Prison Project data are strongly suited to the task, because they include measures of the two most important determinants of imprisonment length – offense severity and criminal history – along with a wide variety of measures related to demographics, lifestyle, and pre-prison work experiences. In addition, we supplemented tests of statistical significance with estimates of effect size to check more carefully for balance on observables (e.g. Connelly et al., 2013), and we performed a variety of sensitivity analyses to increase confidence in the robustness of our findings. That being said, one should always bear in mind the possibility of unobserved confounding variables for the relationship between imprisonment length and employment.

A second limitation concerns potential weaknesses of the data. An advantage of survey data is the ability to collect rich background data, which is essential for a propensity score methodology. Yet a potential downside is that social desirability and memory loss can invalidate responses. We tried to minimize these biases in several ways. For example, during face-to-face interviews we asked retrospective questions about recent events, used

traditional as well as calendar-based questionnaires to measure labor market participation, and acquired additional information on length of imprisonment and criminal history from administrative sources. An important direction for future research is to study the labor market participation of (ex-)prisoners by combining administrative data with self-report data on employment and recidivism.

A third concern is the generalizability of our findings to the wider population of prisoners and to other Western countries. Because of the timing of data collection, short-term prisoners are overrepresented in the current sample. As a result, any findings pertaining to the deteriorating effect of longer imprisonment length are likely to be underestimates. Furthermore, the Netherlands has a relatively mild penal climate, highly restricted access to criminal history records, and a generous social welfare regime. It is therefore a matter of speculation whether we would find similar results using data from other countries, although our findings are most likely to apply to Northern European countries.

Balanced against a concern about generalizability is the paucity of research on the consequences of short prison spells for employment. In the United States and Western Europe, short spells of incarceration are the norm – they are known as prison spells in a European context but jail spells in an American context. Past research on the incarceration-employment relationship, conducted largely in the United States, is limited to prison spells averaging two years. Bearing in mind other differences in the penal and social climates, a study of prison inmates in the Netherlands can fill an important empirical gap concerning the effect of imprisonment length on employment among American jail detainees. Furthermore, given the more humanitarian climate in the Netherlands, we regard the estimates in this study as conservative.

4.11 Conclusion

The present study examines the effect of imprisonment length on employment outcomes in the first six months after release from prison. Because all of the men in this study were incarcerated, we compare groups differing in their imprisonment length, rather than compare men who were incarcerated to men who were not incarcerated. A rich longitudinal dataset comprising 702 pretrial detainees enables us to assess the effect of longer imprisonment on employment outcomes. A variety of post-prison employment variables offers further insight into the labor market prospects of this sample of Dutch ex-prisoners.

The key finding is that, while employment is largely insensitive to imprisonment length among short-term prisoners, there is an apparent threshold at about six months: beyond six months, longer imprisonment corresponds with incremental deterioration in employment prospects. We do note that not all of the sensitivity analyses confirmed this basic pattern.

So while our conclusion should be regarded as provisional, our hope is that follow-up studies will help untangle the impact of imprisonment length on long-term employment prospects (such efforts based on Prison Project data are currently under way). Interestingly, no clear pattern was evident in the intermediate mechanisms considered (prison programming, criminal recidivism), nor was a clear pattern observed with respect to job quality or job stability.

The social context of the Netherlands would seem to indicate that any effects should be conservative relative to short terms of confinement in the U.S. context (specifically, among jail inmates). The fact that employment prospects are apparently worsened among Dutch ex-prisoners serving more than six months suggests that such effects, considered in the context of less generous social welfare and less humanitarian prisons, are likely to be considerably larger.

Appendix 4.A Descriptive statistics for the 55 pre-prison covariates used in the propensity score model (n = 702)

	Mean	SD	Min.	Max.
Demographic Characteristics				
Age	31.85	11.13	18.00	65.00
Non-ethnic Dutch	.31		.00	1.00
Higher level of secondary schooling	.37		.00	1.00
Level of education father				
Low	.33		.00	1.00
High (higher level of secondary schooling)	.21		.00	1.00
Missing	.47		.00	1.00
Level of education mother				
Low	.42		.00	1.00
High (higher level of secondary schooling)	.17		.00	1.00
Missing	.41		.00	1.00
Employment History				
Number of employers	5.91	5.70	.00	20.00
Duration unemployment (years)	3.76	5.45	.00	20.00
Duration longest job (years)	3.71	3.73	.00	13.00
Frequency dismissal	1.36	2.10	.00	8.00
Frequency off-the-books employment	1.75	1.77	.00	5.00
Employment Before Imprisonment				
Non-participant	.22		.00	1.00
Unemployed	.39		.00	1.00
Employee	.26		.00	1.00
Self-employed	.13		.00	1.00
Wage (€)	700.26	1039.74	.00	3200.00
Sources of Income before Imprisonment				
Receive income from others	.12		.00	1.00
Income from off-the-books employment (€)	130.57	306.06	.00	1000.00
Income from illegal act. (€)	667.94	1622.66	.00	6000.00
Income from benefits (€)	329.32	406.76	.00	1100.00
Lifestyle				
Repeated class in school	.29		.00	1.00
Special education	.27		.00	1.00
Ever suspended	.58		.00	1.00
Use of alcohol	1.95	1.69	.00	5.00
Use of drugs	1.70	1.76	.00	4.00
Health	3.66	.89	1.00	5.00
Homeless	.09		.00	1.00
Debts	.63		.00	1.00
Driver's license	.50		.00	1.00
Passport	.47		.00	1.00
ID-document	.63		.00	1.00

Appendix 4.A continued

	Mean	SD	Min.	Max.
Attitude				
Locus of control	2.82	.85	1.00	5.00
Positive attitude towards criminal justice actors	2.80	.42	1.28	4.35
Negative attitude towards law	2.84	.31	1.00	3.90
Motivation to work	3.47	.49	1.00	4.89
Social Bonds				
Children	.46		.00	1.00
Partner	.45		.00	1.00
Criterion Offense Characteristics				
Type of crime				
Violent	.24		.00	1.00
Sex	.02		.00	1.00
Violent property	.16		.00	1.00
Property	.34		.00	1.00
Damage	.08		.00	1.00
Drug offense	.09		.00	1.00
Other/unknown	.07		.00	1.00
Number of crimes in case file	2.71	1.95	1.00	10.00
Maximum penalty (LN)	7.56	.67	4.50	8.90
Pretrial release	.56		.00	1.00
Criminal History				
Number of violent crimes	1.20	1.51	.00	5.00
Number of property crimes	3.82	5.31	.00	18.00
Number of other crimes	2.33	2.39	.00	8.00
Previous prison sentence	.55		.00	1.00
Age of onset	19.39	6.32	11.74	35.20

Appendix 4.B Balance diagnostics, prior to stratification, for the covariates used in the propensity score model (n = 702)

	1-6 weeks		6 wk-3 months		3-4 months		4-6 months		6-12 months			
	n = 132		*n* = 191		*n*= 133		*n* = 127		*n* = 119			
Covariates	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Sig.	R
Demographic Characteristics												
Age	33.02	11.41	30.46	10.45	30.31	10.81	30.69	10.84	31.39	10.81		.04
Non-ethnic Dutch	.23	.43	.34	.48	.35	.48	.31	.46	.27	.45		.01
Higher level of secondary schooling	.45	.50	.34	.47	.41	.49	.35	.48	.34	.47		.05
Level of education father												
Low	.39	.49	.35	.48	.30	.46	.30	.46	.38	.49		.02
High (higher level of secondary schooling)	.24	.43	.12	.33	.19	.39	.20	.41	.20	.40		.01
Missing	.37	.49	.53	.50	.51	.50	.50	.50	.42	.50	*	.01
Level of education mother												
Low	.44	.50	.39	.49	.40	.49	.40	.37	.51	.50		.04
High (higher level of secondary schooling)	.20	.41	.16	.37	.45	.36	.16	.37	.18	.38		.02
Missing	.36	.48	.45	.50	.45	.50	.44	.50	.31	.47		.03
Employment History												
Number of employers	6.15	5.83	6.25	5.92	6.05	6.12	4.90	4.98	6.40	5.82		.02
Duration unemployment (years)	3.98	6.09	3.60	5.30	3.27	5.29	3.24	5.19	4.32	5.64		.00
Duration longest job (years)	3.60	3.95	3.46	3.51	3.41	3.69	3.79	3.95	3.72	3.75		.02
Frequency dismissal	1.33	2.14	1.36	2.09	1.42	2.42	1.16	1.74	1.39	2.04		.00
Frequency off the books employment	1.83	1.82	1.81	1.76	1.71	1.85	1.50	1.64	1.64	1.79		.06
Employment Before Imprisonment												
Non-participant	.27	.44	.23	.42	.20	.40	.23	.42	.23	.42		.02
Unemployed	.40	.49	.36	.48	.41	.49	.39	.49	.39	.49		.01
Employee	.22	.42	.31	.46	.29	.30	.26	.44	.24	.43		.01
Self-employed	.11	.32	.10	.30	.10	.30	.13	.33	.14	.35		.04
Wage (€)	617.91	1016.46	676.84	980.66	677.82	979.69	672.13	1051.95	706.55	1030.51		.02

Appendix 4.B continued

	1-6 weeks		6 wk-3 months		3-4 months		4-6 months		6-12 months			
	n = 132		*n* = 191		*n*= 133		*n* = 127		*n* = 119			
Covariates	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Sig.	R
Sources Of Income Before Imprisonment												
Receive income from others	.14	.34	.15	.36	.16	.37	.10	.30	.12	.32		.03
Income from off the books employment (€)	146.36	334.00	163.80	344.07	119.58	297.15	77.59	240.54	180.63	359.02		.01
Income from illegal act. (€)	613.78	1,600.66	607.93	1,507.40	952.85	1,875.71	590.15	1,445.26	862.45	1,884.20		.04
Income from benefits (€)	369.94	418.05	302.13	418.27	287.99	394.80	317.03	403.78	321.23	407.49		.03
Lifestyle												
Repeated class in school	.32	.47	.37	.48	.28	.45	.26	.44	.40	.49		.01
Special education	.25	.44	.29	.45	.26	.44	.25	.44	.27	.45		.00
Ever suspended	.52	.50	.57	.50	.58	.50	.50	.50	.69	.47	*	.07
Use of alcohol	2.11	1.81	2.00	1.72	1.83	1.61	2.09	1.66	1.97	1.59		.02
Use of drugs	1.72	1.72	1.75	1.76	1.82	1.80	1.50	1.72	1.74	1.76		.02
Health	3.58	.84	3.77	.90	3.70	1.01	3.75	.88	3.62	.88		.00
Homeless	.14	.34	.09	.29	.05	.22	.13	.33	.10	.30		.01
Debts	.67	.47	.60	.49	.57	.50	.61	.49	.62	.49		.02
Driver's license	.50	.50	.48	.50	.46	.50	.50	.50	.48	.50		.00
Passport	.42	.50	.46	.50	.50	.50	.51	.50	.45	.50		.03
ID-document	.64	.48	.60	.49	.66	.48	.59	.49	.59	.49		.03
Attitude												
Locus of control	2.82	.72	2.69	.82	2.73	.84	2.86	.89	2.87	.82		.05
Positive attitude towards criminal justice actors	2.74	.44	2.83	.40	2.79	.48	2.81	.44	2.82	.44		.04
Negative attitude towards law	2.86	.31	2.81	.32	2.80	.32	2.84	.37	2.86	.28		.02
Motivation to work	3.45	.49	3.50	.47	3.48	.54	3.49	.48	3.53	.52		.04

Appendix 4.B continued

	1-6 weeks		6 wk-3 months		3-4 months		4-6 months		6-12 months			
	n = 132		*n* = 191		*n*= 133		*n* = 127		*n* = 119			
Covariates	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Sig.	R
Social Bonds												
Children	.39	.49	.38	.49	.46	.50	.47	.50	.49	.50		.08
Partner	.39	.49	.40	.49	.47	.50	.43	.50	.51	.50		.08
Criterion Offense Characteristics												
Type of crime												
Violent	.27	.44	.23	.42	.21	.41	.28	.45	.20	.40		.02
Sex	.02	.12	.02	.13	.04	.19	.02	.13	.03	.16		.02
Violent property	.10	.30	.13	.34	.17	.37	.17	.38	.29	.46	***	.16
Property	.43	.50	.41	.49	.35	.48	.29	.46	.22	.42	**	.16
Damage	.09	.29	.06	.24	.08	.27	.06	.23	.04	.20		.05
Drug offense	.08	.28	.09	.29	.11	.31	.10	.30	.13	.33		.04
Other/unknown	.02	.12	.06	.24	.05	.22	.08	.27	.09	.29		.09
Number of crimes in case file	2.23	1.57	2.31	1.38	2.44	1.46	2.89	1.92	3.13	2.07	***	.19
Maximum penalty (LN)	7.47	.52	7.52	.64	7.58	.72	7.61	.73	7.84	.60	***	.17
Pretrial release	.72	.45	.65	.48	.48	.50	.32	.47	.24	.43	***	.36
Criminal History												
Number of violent crimes	1.09	1.51	1.28	1.56	1.14	1.45	1.28	1.52	1.24	1.53		.02
Number of property crimes	4.03	5.54	3.86	5.33	3.53	5.30	1.91	4.55	3.05	4.29		.07
Number of other crimes	2.79	2.64	2.43	2.53	2.05	2.37	1.91	2.11	2.43	2.43	*	.07
Previous prison sentence	.62	.49	.55	.50	.52	.50	.50	.50	.50	.50		.06
Age of onset	19.24	5.95	18.14	5.23	18.81	6.09	19.73	6.61	5.52	5.52		.03

NOTE: For all covariates, p values reflect one-way analysis of variance (ANOVA) tests for the equality of means.
ABBREVIATIONS: SD= standard deviation (omitted for dummy variables)
* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed tests).

Appendix 4.C Balance diagnostics, after stratification, for the covariates used in the propensity score model (n = 666)

Covariates	1-6 weeks *n* = 125	6 wks-3 months *n* = 190	3-4 months *n* = 131	4-6 months *n* = 118	6-12 months *n* = 102	Significance			
	Mean	Mean	Mean	Mean	Mean	Length	Length* quintile	Both	R
Demographic Characteristics									
Age	33.59	30.21	30.31	31.33	33.61				.01
Non-ethnic Dutch	.20	.38	.34	.24	.33		*	*	.02
Completed higher level of secondary schooling	.46	.31	.44	.36	.38				.03
Level of education father									
Low	.39	.33	.29	.32	.33				04
High (higher level of secondary schooling)	.18	.14	.21	.18	.35	*			.12
Missing	.42	.53	.50	.50	.32				.07
Level of education mother									
Low	.59	.34	.39	.36	.55	*			.02
High (higher level of secondary schooling)	.13	.20	.15	.18	.20				.03
Missing	.28	.45	.46	.46	.26				.01
Employment History									
Number of employers	5.50	6.02	6.21	4.96	6.52				.02
Duration unemployment (years)	4.52	3.62	3.19	3.31	4.05				.03
Duration longest job (years)	4.07	3.43	3.39	3.62	3.97				.00
Frequency dismissal	1.48	1.37	1.49	1.06	1.42				.03
Frequency off the books employment	1.82	1.73	1.72	1.56	1.80				.02
Employment Before Imprisonment									
Non-participant	.20	.22	.21	.24	.24				.03
Unemployed	.44	.35	.40	.38	.39				.01

Appendix 4.C continued

Covariates	1-6 weeks $n = 125$	6 wks-3 months $n = 190$	3-4 months $n = 131$	4-6 months $n = 118$	6-12 months $n = 102$				
	Mean	Mean	Mean	Mean	Mean	Length	Significance Length* quintile	Both	R
Employee	.18	.32	.30	.25	.24				.01
Self-employed	.18	.11	.09	.12	.12				.04
Wage (€)	646.11	691.83	653.87	676.20	632.74				.00
Sources Of Income Before Imprisonment									
Receive income from others	.09	.14	.15	.10	.11				.00
Income from off the books employment (€)	136.65	156.78	110.99	65.41	197.01				.00
Income from illegal act. (€)	530.52	629.45	926.08	600.73	704.37				.03
Income from benefits (€)	490.00	301.83	298.10	356.13	324.59				.08
Lifestyle									
Repeated class in school	.23	.37	.27	.24	.28				.02
Special education	.29	.31	.27	.29	.30				.00
Ever suspended	.57	.60	.58	.48	.66				.00
Use of alcohol	1.71	1.91	1.83	1.99	2.47				.12
Use of drugs	1.81	1.70	1.75	1.48	1.72				.03
Health	3.48	3.79	3.74	3.73	3.52				.00
Homeless	.11	.09	.04	.17	.07				.00
Debts	.76	.58	.56	.62	.63				.06
Driver's license	.49	.49	.45	.48	.50				.00
Passport	.46	.49	.49	.47	.41				.03
ID-document	.64	.63	.67	.67	.57				.02

Appendix 4.C continued

Covariates	1-6 weeks $n = 125$	6 wks-3 months $n = 190$	3-4 months $n = 131$	4-6 months $n = 118$	6-12 months $n = 102$	Significance			
	Mean	Mean	Mean	Mean	Mean	Length	Length* quintile	Both	R
Attitude									
Locus of control	2.93	2.68	2.72	2.84	2.90				.02
Positive attitude towards criminal justice actors	2.74	2.83	2.78	2.82	2.84				.05
Negative attitude towards law	2.93	2.81	2.79	2.85	2.84				.05
Motivation to work	3.37	3.50	3.45	3.49	3.47				.05
Social Bonds									
Children	.54	.41	.45	.46	.40				.05
Partner	.54	.45	.46	.39	.37				.11
Criterion Offense Characteristics									
Type of crime									
Violent	.25	.23	.21	.20	.24				.02
Sex	.10	.01	.05	.01	.01	**	***	***	.12
Violent property	.13	.16	.16	.10	.19				.02
Property	.27	.36	.36	.35	.37		*		.05
Damage	.18	.05	.07	.06	.07		*		.09
Drug offense	.05	.10	.10	.09	.09				.03
Other/unknown	.02	.08	.06	.09	.02				.01
Number of crimes in case file	2.91	2.42	2.39	2.54	2.49				.05
Maximum penalty (LN)	7.56	7.55	7.57	7.40	7.65				.00
Pretrial release	.70	.58	.51	.52	.51				.12

Appendix 4.C continued

Covariates	1-6 weeks $n = 125$	6 wks-3 months $n = 190$	3-4 months $n = 131$	4-6 months $n = 118$	6-12 months $n = 102$	Significance			
	Mean	Mean	Mean	Mean	Mean	Length	Length* quintile	Both	R
Criminal History									
Number of violent crimes	1.15	1.31	1.07	1.25	1.23				.00
Number of property crimes	3.07	3.63	3.42	3.87	5.06				.10
Number of other crimes	2.31	2.30	2.02	2.24	2.65				.03
Previous prison sentence	.50	.55	.50	.52	.73				.10
Age of onset	22.29	18.32	18.92	19.25	19.92	**	**	*	.07

NOTE: For all covariates, p values reflect one-way analysis of variance (ANOVA) tests for the equality of the means.
* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed tests).

Appendix 4.D Propensity score quintile conditional on imprisonment length (n = 666)

	Observed dose of imprisonment				
Predicted scores	1-6 weeks	6 wks-3 months	3- 4 months	4-6 months	6-12 months
1st quintile	.44	.24	.15	.08	.03
2nd quintile	.32	.28	.17	.09	.06
3rd quintile	.12	.24	.26	.21	.14
4th quintile	.10	.16	.27	.25	.25
5th quintile	.02	.07	.15	.37	.52
N	125	190	131	118	102

Appendix 4.E Details on estimation of a generalized propensity score with a continuous treatment

In addition to estimating a propensity score model for an ordered treatment, a propensity score model for a continuous treatment was also considered. For this supplementary analysis, we retained imprisonment length in its original metric, representing the number of days of prison confinement. The rationale of the approach is developed by Hirano and Imbens (2004) and Imai and Van Dyck (2004), with an application available in Bia and Mattei (2008). In this appendix, we provide a detailed description of the method, but we begin in the next paragraph with a very brief, non-technical overview.

In the current analysis, the generalized propensity score (GPS) represents the estimated probability of the residual from a log-linear regression of imprisonment length on all covariates. This probability derives from the standard normal or *z*-distribution, and therefore assumes that the residuals are normally distributed (an assumption that can be empirically verified). Following estimation of the GPS but prior to estimation of the dose-response function, it is important to ascertain that subjects with different assigned imprisonment lengths are "balanced" with respect to the pre-prison covariates. In other words, conditional on the GPS, there should be no systematic tendency for subjects possessing different imprisonment length "doses" to differ with respect to the covariates. A procedure which involves stratifying on quantiles of imprisonment length (and then stratifying further on quantiles of the GPS) provides a means of testing the balancing property of the model. Estimation of the dose-response function proceeds after the balancing property and the support condition have been confidently established. A GPS-adjusted probability of employment for a given imprisonment length dose can be calculated, which averages over each subject's dose-specific GPS. This can be performed for each imprisonment length of interest, which in the present study is 5 days to 415 days, and then summarized in a graph.

In more technical terms, estimation of the dose-response function using a generalized propensity score consists of three basic steps. The first step is estimation and diagnosis of the GPS. We performed maximum likelihood (ML) regression of the natural logarithm of imprisonment length (denoted T_i, for "treatment dosage") on all covariates, along with the squared and interaction terms necessary to maximize balance on the covariates. The model is represented straightforwardly as follows:

$$ln(T_i) = \beta_0 + \sum_{j=1}^{K} \beta_j X_{ij} + \varepsilon_i$$

where $i = 1,\ldots,N$ indexes subjects and $j = 1,\ldots,K$ indexes regressors. Following estimation, the normality of the residuals was confirmed from the non-parametric, Kolmogorov-Smirnov test. These residuals are shown in Appendix F.

We then evaluated each of the residuals with respect to the standard normal probability density function:

$$P_i = \phi\left(\frac{ln(T_i) - X\beta_i}{\sigma_\varepsilon}\right)$$

where $X\beta_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_i X_{iK}$ references the linear predictor for subject i (obtained using the estimates from the ML regression model above), σ_ε is the square root of the ML estimate of the model variance, and $\phi(\cdot)$ is the standard normal density evaluated at the argument (i.e., the height of the standard normal distribution at the evaluation point). By construction, P_i is the GPS, formally defined as the conditional density of treatment given the covariates (Hirano and Imbens, 2004). Less formally, the GPS is just the probability assigned to a z-score, where z_i is defined as $ln(T_i)$ with reference to the mean ($X\beta_i$) and standard deviation (σ_ε) of a normal random variable.

The second step is evaluation of the balancing property of the GPS. There are several ways to do so, but we relied on the method proposed by Hirano and Imbens (2004), which involves blocking on imprisonment length and the estimated GPS. After estimation of the GPS from the first step, imprisonment length is divided into four equal-sized strata (quartiles). Then, within each stratum, an auxiliary GPS is calculated by evaluating each subject's linear predictor with respect to the median imprisonment length for the stratum:

$$P_i^r = \phi\left(\frac{Median\left(ln\left[T_{i\in r}\right]\right) - X\beta_i}{\sigma_\varepsilon}\right)$$

where $r = 1,\ldots,S$ indexes imprisonment length strata and the linear predictor and standard deviation are the same terms obtained from the regression model in the first step. The outcome of this step is the creation of four such auxiliary variables for each subject – one for the evaluation with respect to the median of each imprisonment length stratum. To evaluate covariate balance, each auxiliary GPS is then divided into five equal-sized blocks (quintiles). Balance is tested by computing mean differences of each covariate between subjects assigned to the same GPS block but classified into different imprisonment length strata:

$$t_j^{g,l} = \frac{\mu_{j,1\in r}^{g,l} - \mu_{j,2\notin r}^{g,l}}{\sqrt{\frac{\left(N_{1\in r}^{g,l} - 1\right)s_{j,1\in r}^{g,l} + \left(N_{2\notin r}^{g,l} - 1\right)s_{j,2\notin r}^{g,l}}{N_{1\in r}^{g,l} + N_{2\notin r}^{g,l} - 2}}\sqrt{\frac{1}{N_{1\in r}^{g,l}} + \frac{1}{N_{2\notin r}^{g,l}}}}$$

where $g = 1,\ldots,H$ indexes auxiliary GPS's, $l = 1,\ldots,M$ indexes blocks of the auxiliary GPS and, as before, $j = 1,\ldots,K$ indexes regressors and $r = 1,\ldots,S$ indexes imprisonment length strata. Note that this yields an independent-samples t-test which is specific to block l of auxiliary GPS g. Because they are independent across blocks, the means and variances from these t-tests can be weighted and combined to yield a single overall test of balance of subjects in imprisonment length stratum r relative to all other subjects.

The foregoing procedure is repeated for each imprisonment length stratum. Of the 55 covariates tested in four such comparisons – resulting in 220 total groupwise comparisons – just 4 covariates are statistically significant ($p < .05$, two tails). This indicates that the covariates are strongly balanced by the GPS. For reference, prior to conditioning on the GPS, 12 of the 55 covariates are significantly different in at least one comparison, yielding 20 of the 220 total groupwise comparisons that are significant.

The third and final step is evaluation of common support and estimation of the dose-response function. We first plotted the auxiliary GPS's estimated from the second step, separately for the subjects assigned to a given imprisonment length stratum and the subjects who were not assigned to the stratum. Inspection of Appendix G reveals that the distributions overlap to a degree that we have confidence the support condition is satisfied. We then regressed the employment outcomes on imprisonment length and the GPS (as well as the product of the two). Because the key response variables considered here are binary measures of employment, a pair of logistic regression models was specified:

$$ln\left[\frac{Pr(Y_i = 1)}{1 - Pr(Y_i = 1)}\right] = \delta_0 + \delta_1 T_i + \delta_2 P_i + \delta_3 T_i \times P_i$$

where T_i is imprisonment length and P_i is the GPS estimated from the first step. To evaluate sensitivity of the dose-response estimates, several polynomial functions of imprisonment length and the GPS were considered (e.g., quadratic and cubic functions and their interactions), although we limit our attention here to the simpler, linear functional form with the interaction. Note that the coefficients obtained from this model have no meaningful interpretation and are instead required for estimation of the dose-response function.

After obtaining the results from the logistic regression model, it is finally possible to estimate the dose-response function, or the GPS-adjusted probability of employment for a given imprisonment length. Doing so requires first calculating a dose-specific GPS that evaluates each subject's linear predictor (from the first step) with respect to a specified imprisonment length:

$$P_i^t = \phi\left(\frac{ln(T_i = t) - X\beta_i}{\sigma_\varepsilon}\right)$$

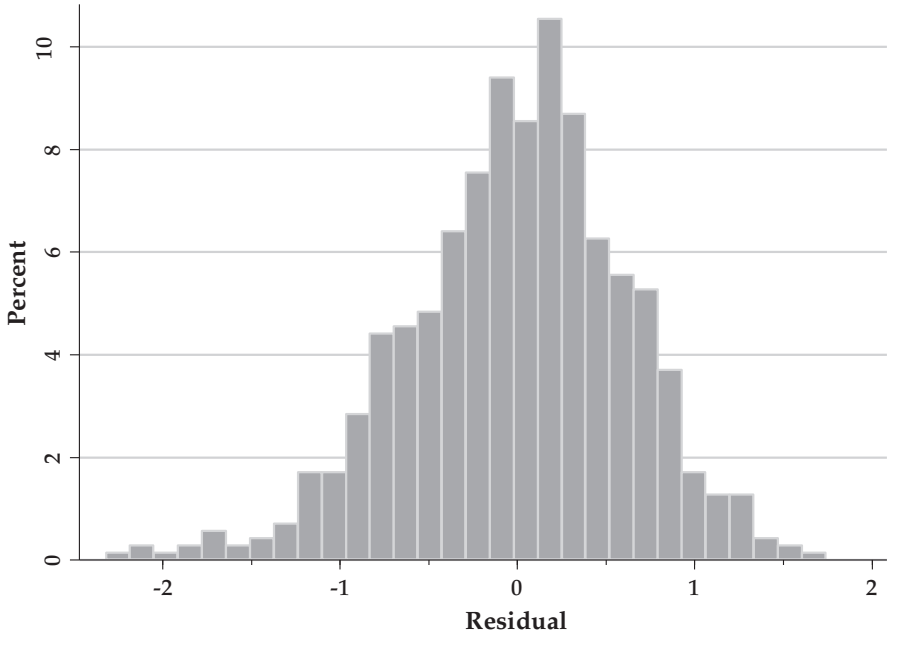
where $T_i = t$ denotes the treatment dosage of interest. We can then use the parameter estimates from the logistic regression model to compute a predicted probability of employment that fixes $T_i = t$ for each subject and then averages over the dose-specific GPS's for all subjects:

$$\overline{Pr\left(Y_i = 1 \middle| T_i = t\right)} = \frac{1}{N}\sum_{i=1}^{N} \frac{exp\left[\hat{\delta}_0 + \hat{\delta}_1\left(T_i = t\right) + \hat{\delta}_2 P_i^t + \hat{\delta}_3\left(T_i = t\right) \times P_i^t\right]}{1 + exp\left[\hat{\delta}_0 + \hat{\delta}_1\left(T_i = t\right) + \hat{\delta}_2 P_i^t + \hat{\delta}_3\left(T_i = t\right) \times P_i^t\right]}$$

where the coefficients are the estimates obtained from the logistic regression model described above. In principle, there are as many GPS-adjusted response probabilities as there are treatment doses, and in the present study, imprisonment length varies from 5 days to 415 days (yielding up to 411 imprisonment length doses for each subject). Standard errors for the average predicted probabilities are obtained by the bootstrap (with 500 replications), which accounts for the uncertainty introduced by the coefficients and the GPS, both sets of which are themselves estimates of unknown quantities.

The key findings are provided in the two graphs shown in Figure 4.4 in the main text. Using the procedures outlined above, the graphs provide the mean probability of employment for specific imprisonment length "doses," conditional on the GPS (and thus the covariates indexed by it). Note that, because each subject possesses a "potential outcome" under each imprisonment length dose, all subjects contribute to the estimates of the dose-specific means and standard errors. This means that the means and confidence intervals shown in the graph are produced from 288,522 predicted employment probabilities (702 subjects × 411 imprisonment length doses).

Appendix 4.F Distribution of the residual of logged imprisonment length, following estimation of the generalized propensity score model



Appendix 4.G Common support distributions following estimation of the generalized propensity score model (n = 702)