Universidade Federal do Ceará

02 | CAEN PÓS-GRADUAÇÃO EM ECONOMIA

SÉRIE
ESTUDOS ECONÔMICOS CAEN

Job Search, Conditional Treatment and Recidivism: The Employment Services for Ex-Offenders Program Reconsidered

Herman J.Bierens
José Raimundo Carvalho

FORTALEZA • ABRIL • 2007

Job Search, Conditional Treatment and Recidivism: The Employment Services for Ex-Offenders Program Reconsidered

Herman J. Bierens*
Department of Economics, Pennsylvania State University

José R. Carvalho†
CAEN, Universidade Federal do Ceará, Brazil.

April 13, 2007

Abstract

The objective of this paper is to re-evaluate the effect of the 1985 "Employment Services for Ex-Offenders" (ESEO) program on recidivism, in San Diego, Chicago and Boston. The initial group of program participants was split randomly in a control group and a treatment group. The actual treatment (mainly being job related counseling) only takes place conditional on finding a job, and not having been arrested, for those selected in the treatment group. We use interval-censored proportional hazard models for job search and recidivism time, where the latter model incorporates the conditional treatment effect, depending on covariates. We find that the effect of the program depends on location and age. The ESEO program reduces the risk of recidivism only for ex-inmates over the age of 27 in San Diego and Chicago, and over the age of 36 in Boston, but increases the risk of recidivism for the other ex-inmates in the treatment group.

*Correspondence address: Herman J. Bierens, Department of Economics, Pennsylvania State University, 608 Kern Graduate Building, University Park, PA 16802. E-mail: hbierens@psu.edu.

†The financial support of CAPES Foundation, Brazil, is gratefully acknowledged.

Key words: Job search, Recidivism, Conditional treatment, Program evaluation, Duration models, Interval-censoring.

1 Introduction

During the period of $1980 - 1985$, the National Institute of Justice (NIJ) sponsored a controlled experiment to evaluate the impact of reemployment programs for recent released prisoners. Three well established programs were chosen, COERS in Boston, JOVE in San Diego and Safer Foundation in Chicago, to participate in the Employment Services for Ex-Offenders Program, henceforth ESEO. A total of $2,045$ prisoners who voluntarily accepted to participate were randomly assigned to either a treatment group or a control group. Those in the first group received, besides the normal services (orientation, screening, evaluation, support services, job development seminar, and job search coaching), special services which consisted of an assignment to a follow-up specialist who provided support during the job search and the 180 days following job placement. The control group received only normal services. The inclusion of special services was a major response to the increasing belief that some past employment programs had failed because ex inmates lost contact with their original programs.

Using OLS regressions, Milkman (2001) found that the effect of the special services program is negligible. However, this evaluation of the ESEO program did not account for the conditional feature of the treatment. The timing of the treatment was completely neglected and, as will be shown below, this is a very important characteristic of the program under evaluation.

The objective of our paper is to re-evaluate the effect of the ESEO program on recidivism using duration models for job search and recidivism (the latter being defined as the time between release and the first arrest), with the conditional treatment incorporated in the latter duration model. Both durations are interval censored.

Carvalho and Bierens (2007) have studied the effect of the ESEO program on recidivism using a bivariate mixed proportional Weibull hazard model for job search and recidivism with common heterogeneity, where both durations are assumed to be independent conditional on the covariates and the heterogeneity variable. The latter is assumed to be Gamma distributed, and integrated out to make the two durations dependent conditional on the covariates only. However, the problem with this approach is that then the job search and recidivism durations are positively related: The longer the job search takes, the longer the time between release and rearrest. This is clearly unrealistic.

Therefore, in the current paper we take a completely different approach, by incorporating treatment in the conditional distribution of the recidivism duration given the covariates **and** the job search duration. This approach is inspired by (but not exactly equal to) the approach of Abbring and Van den Berg (2003), in particular their Model 1a. These authors propose a bivariate mixed proportional hazard model, where one duration is the timing of the treatment and the other one is the duration of interest that is affected by the treatment. In the case of the ESEO program the duration that triggers the treatment is the job search, and the duration of interest is recidivism. However, since both durations are interval-censored and all the covariates are discrete we cannot take unobserved heterogeneity into account, because it is shown in Bierens (2007) that then the mixed proportional hazard model is not identified. Moreover, Abbring and Van den Berg (2003) do not take the presence of a control group into account.

The setup of the paper is as follows. In Section 2 we discuss the ESEO program in more detail, and in Section 3 we describe the data set and the variables involved. Section 4 deals with the attrition problem, which is substantial. It appears that attrition is affected by the group assignment: the attrition rate is higher for the control group than for the treatment group. In Section 5 we conduct a preliminary data analysis by estimating and comparing the interval probabilities for both durations without using covariates, in order to check for evidence of dependence of recidivism on job search, and evidence of a treatment effect. We find neither, and therefore we conclude that at least for the control group the job search and recidivism durations are independent conditional on the covariates, and that if there is a treatment effect then it will likely work via covariates. Section 6 deals with model specification and estimation. In Section 6.1 we discus the Abbring and Van den Berg (2003) model, and its relation to our model. In Section 6.2 we estimate and test an interval-censored proportional hazard model for the job search, where in first instance the integrated hazard is left unspecified. Only the location (Chicago, San Diego or Boston) seems to matter for the job search duration, and the results indicate that a Weibull baseline hazard is appropriate for job search. Similarly, in Section 6.3 we estimate and test an interval censored proportional hazard model for the recidivism duration of the control group only. It appears that none of the covariates matter and that the baseline hazard is constant, hence the distribution involved is exponential. In Section 6.4 we extend this exponential model for recidivism to a model that incorporates treatment via covariates, depending on the job search duration, and merge it with the job search model, which are then estimated jointly by maximum likelihood. The results are presented in Section 6.5. It appears that a treatment effect is present, but its magnitude and direction depends on the location and on age. In Section 6.6 we compare the results with the preliminary data analysis, and explain why we did not find evidence of dependence of the job search and recidivism durations and the presence of a treatment effect. In Section 6.7 we compute the actual treatment effect, in terms of expected increase or reduction of the recidivism duration in months, on the basis of our estimations results. Finally, in Section 7 we explain and discus the results, and make some concluding remarks.

All computations have been conducted with the free software package *EasyReg International* developed by the first author. See Bierens (2006).

2 The ESEO program

During the last decades, sociologists and economists have been devising programs to ease the difficult transition faced by ex-offenders during the period of time between release and reintegration into society. As experience has accumulated, a fundamental goal to a complete reintegration turned out to be job placement. A good job would be necessary not only to provide the basic needs for survival in the short run but also as a key element to secure self-esteem, security and sense of integration in the society as whole. Hence, sociological and economic theory have provided enough justification for the existence of employment services programs for ex-offenders.

The Life Insurance for Ex-offenders (LIFE) and the Transitional Aid for Ex-offenders (TARP) are two early examples of employment services for ex-offenders. Both programs offered financial assistance as well as job placement services. The two programs reached similar conclusions: while financial assistance appeared to decrease the recidivism rate, job placement had little or no effect on reducing criminal activity, unless for those who succeeded in securing a job for a long time. These early results should not be interpreted as a failure but, in fact, should be viewed as just a first step to the design of better programs. The lack of follow-up after placement was conjectured by Milkman et al. (1985) as the main obstacle to the complete success of such programs.

The new paradigm of employment services for ex-offenders have resulted in the appearance of programs that had a strong preoccupation with the post-placement of their clients. These programs have designed follow-up strategies to overcome the major criticism of past experience. Among various programs, three deserve recognition for both being successful and having similar structures: the Comprehensive Offender Resource System, in Boston, the Safer Foundation, in Chicago and Project JOVE, in San Diego. Not surprisingly, the US. Department of Justice saw this as a opportunity for assessing the efficacy of employment services programs that contained a follow-up component. Then, in 1985 the Department of Justice funded a research performed by the Lazar Institute from McLean, VA.

In the ESEO program, after being assigned to either the control or treatment group, the clients step inside the intake unit, where they received initial orientation, screening and evaluation by an intake counsellor. While still in this first phase, to secure survival up to the job search phase, the intake counsellor offered minimal assistance services such as food, transportation, clothing and etc.

After intake, the client enters the second phase that will prepare him/her to develop job search skills: brief job development seminar which deals with issues like appropriate dress and deportment, typical job rules, goal setting, interviewing techniques, and job hunt strategy. It is assumed that the time spent in the first and second phases are not random and negligible compared to the search phase and the average duration of the outcome. The next and final phase before possible treatment is the job search assistance. This is the traditional job search assistance type of service, as described by Heckman et al. (1999). The job search assistance in the ESEO program is offered equally to both control and treatment groups.

The actual treatment starts upon the first job placement. The people in the control group did not receive help after placement, whereas the people in the treatment group received follow-up help. The follow-up special services consisted mainly of crisis intervention, counselling and, whenever necessary, re-employment assistance, during a period of six month after the first placement. See Timrots (1985), Milkman et al. (1985) and Milkman (2001) for the details of the programs and treatment.

Of course, the ESEO program may not be identical in the three locations. Therefore, possible heterogeneity will be controlled for via location dummy variables.

3 The ESEO data set

The ESEO data set consist of $2,045$ individuals who participated in one of the three programs: 511 in Boston, 934 in Chicago and 600 in San Diego. However, the ICPSR[1] only made available 1074 usable observations: 325 in Boston, 489 in Chicago and 260 in San Diego. A large amount of information, sometimes very detailed, was collected from all three sites.

A first important empirical issue is related to the characterization of the population being sampled. Unless very special assumptions are made, the validity of our findings can not be extrapolated beyond the population under sampling. In order to be eligible to participate in the ESEO program an individual must have the following background:

1. Participants voluntarily accepted program services;
2. Participants had been incarcerated at an adult Federal, State, or local correctional facility for at least 3 months and had been released within 6 months of program participation;
3. Participants exhibited a pattern of income-producing offenses.

From the eligibility criteria it is clear that our population is a special, indeed a very special, subset of the population of ex-offenders. Also, since participation is voluntary and there is no information on non-participants (those who did not choose to participate even though they fulfilled requirements 2 and 3.), it is not possible to assess the potential bias induced by this selection scheme.

Given the initial sample, the individuals were randomly assigned to either the treatment or control group. Controls receive the standard services and treatments received, in addition to that, emotional support and advocacy during the follow-up period of 180 days after placement. Two durations are of great importance, time spent searching a job and recidivism time. These two variables are interval-censored, however: They are only observed in the form of intervals.

The point of departure for the choice of the covariates of the recidivism duration model is Schmidt and Witte (1988): age at release, time served for

[1]The data set used in our present analysis comes from the Inter-University Consortium for Political and Social Research, henceforth ICPSR, University of Michigan, under the study number 8619.

the sample sentence, sex, education, marital status, race, drug use, supervision status, and dummies that characterize the type of recidivism. However, we have also paid close attention to the criminologic literature in recidivism, for instance Gendreau et al. (1996).

The literature on unemployment (and job search) duration has been refined since the 70's. Nowadays, it has a status of a complete theory of unemployment, as it appears in Pissarides (2000). Its empirical contents has been developed since the late 70's and this first wave of empiricism is characterized for being concerned with "reduced" type models. A good account of this first phase can be found in Devine and Kiefer (1991). A final wave is characterized by advocating a "structural" approach to estimation and inference in such models. An updated account of that appears in Van den Berg (1999, 2000). There has been also studies close to ours that try to measure the effect of programs in a context of a model of unemployment and job search duration. See for instance, Abbring and Van den Berg (2003), Abbring et al. (2005), Eberwein et al. (1997) and Van den Berg et al. (2004). In view of those studies, the following set of covariates has been singled out, next to the group assignment indicator:

- $G = 1$ if selected in the treatment group, $= 0$ if selected in the control group
- $DRUGS$: Indicator for drugs use during the last 5 years;
- $WHITE$: Race indicator;
- $MALE$: Gender indicator;
- $EDUC\_L = 1$ if years of schooling ≤ 8, $= 0$ otherwise;
- $EDUC\_H = 1$ if years of schooling > 12, $= 0$ otherwise;
- AGE: Age in years;
- $AGE1ARR$: Age of first arrest, in years;
- $SANDIEGO$: Indicator for San Diego;
- $CHICAGO$: Indicator for Chicago;
- $BOSTON = 1 - SANDIEGO - CHICAGO$

The endogenous variables are:

- A: Indicator for attrition. $A = 1$ means the individual is either a "no show" or a "drop-out", $A = 0$ otherwise;

- T_s ($s = search$) is the duration of the job search, i.e., the time between the date of release and the date of placement in the first job after release;

- T_c ($c = crime$) is the recidivism time, i.e., the time between the date of release and the date of the first arrest after release.

As to attrition, one could argue that "no show" and "drop–out" deserve separate treatment, as pointed out by Bloom (1984) and Heckman et al. (1998). However, our data set does not allow for this disaggregation. Besides, most drop-outs occur very early in the program.

The job search duration does not need any explanation, but the meaning of "recidivism" is not unambiguous. There are two ways to measure recidivism outcomes in the ESEO program, via count data or duration data. Detailed data on the number of arrests from date of released to the end of the program for all clients was gathered in the respective state police departments. That was the data used in the original evaluation made by Milkman et al. (1985). Also, data of the first arrest after release is available. In the criminology literature three possible definitions of recidivism are considered: rearrest, reconviction and reincarceration. It seems that rearrest has been proven to be the most reliable among the three possible measures, as reported in Beck and Shipley (1989) and Maltz (1984). The latter is what we will use as the duration of recidivism. Thus, in the sequel "(duration of) recidivism" should be interpreted as the time between release and first rearrest.

Both durations are interval-censored: $T_s \in (a, b]$, $T_c \in (p, q]$, $b \leq p$, where $(a, b], (p, q] \in \{(0, 1], (1, 6], (6, 12], (12, \infty)\}$. The unit of measurement for these durations is months. These events are only observed if $A = 0$.

As said before, the ex-convicts in the sample have been randomly assigned to either a control group ($G = 0$) or a treatment group ($G = 1$). Both groups get standard assistance with the job search. Treatment consists of extra help after finding a job, and is therefore conditional on the job search duration T_s. The purpose of this study is to determine whether this treatment has an effect on the risk of recidivism, given the covariates listed above.

4 Attrition

There is a substantial number of individuals in the sample who do not show up for the common part of the program, i.e., the assistance with the job search. With a few exceptions, attrition occurs straight away after release. It is therefore reasonable to assume that, conditional on the covariates (stacked in a vector X), A is independent of T_s and T_c, because in most cases the attrition decision $A = 1$ is made before T_s or T_c are realized.

The conditional attrition probability $P[A = 1|X]$ has been modeled as a Logit model. It is conceivable that attrition is also affected by the group assignment: The prospect of receiving treatment ($G = 1$) may lead to a lower attrition probability. Therefore, next to the original covariates we have included the group assignment G and the products of G with each of the covariates in the Logit model for A. The initial Logit model has been subjected to a sequence of Wald and likelihood ratio tests to clean the model of insignificant covariates. The details of this cleaning process can be found in the Appendix; only the final results are presented here, in Table 1.

It follows from Table 1 that males have a higher attrition rate than females, but only if selected in the control group. Moreover, the attrition rates in Chicago are (significantly) higher for the control group than for the treatment group, and the same applies to San Diego.

Table 1. Logit results for attrition

Covar.$\times(1-G)$	Estimates	t-val.	Covar.$\times G$	Estimates	t-val.
$MALE$	0.7835337	3.93	$CHICAGO$	0.5250147	2.71
$CHICAGO$	1.7985664	7.94	$SANDIEGO$	0.4278509	2.05
$SANDIEGO$	1.0452205	3.55	1 (not $\times G$)	-0.7661767	-5.26
Log-likelihood:	-654.831		Sample size:	1074	

To compare the attrition rates in San Diego and Chicago with the attrition rate in Boston, we have to compare the coefficients of the location dummies with the intercept, which leads to the conclusion that the attrition rates in Chicago and San Diego are much higher than in Boston, in particular for the control group, and within the control group the attrition rate in Chicago is higher than in San Diego, although for the treatment group the attrition rates in Chicago and San Diego are not significantly different (the p-value of the Wald test involved is 0.62037). The reason may be that the standard service in assisting the ex-inmates with the job search is much more effective in Boston than in San Diego and Chicago, and that therefore the motivation for participation in the program is higher in Boston than in the other two cities. This explanation appears to be corroborated by the results below for the job search duration.

Finally, note that the dependence of attrition on the group assignment does not conflict with our assumption that conditional on the covariates T_s and T_c are independent of A, because G is determined randomly.

5 Preliminary data analysis

As said before, the durations T_c and T_s are only observed in the form of interval indicators, for the intervals $(0,1], (1,6], (6,12]$ and $(12,\infty)$. Table 2 below presents the number of observations in each interval and combination of intervals, for both groups as well as separately for the treatment group ($G=1$) and the control group ($G=0$).

Table 2. Observations per interval (A=0)

$T_s \diagdown T_c$	$(0,1]$	$(1,6]$	$(6,12]$	$(12,\infty)$	*Total*
$(0,1]$	12	56	43	152	263
$(1,6]$	6	44	39	112	201
$(6,12]$	1	7	6	20	34
$(12,\infty)$	0	1	1	3	5
Total	19	108	89	287	503

Treatment group only:

$T_s \diagdown T_c$	$(0,1]$	$(1,6]$	$(6,12]$	$(12,\infty)$	*Total*
$(0,1]$	9	39	32	105	185
$(1,6]$	5	35	30	81	151
$(6,12]$	1	4	5	18	28
$(12,\infty)$	0	1	1	3	5
Total	15	79	68	207	369

Control group only:

$T_s \diagdown T_c$	$(0,1]$	$(1,6]$	$(6,12]$	$(12,\infty)$	*Total*
$(0,1]$	3	17	11	47	78
$(1,6]$	1	9	9	31	50
$(6,12]$	0	3	1	2	6
$(12,\infty)$	0	0	0	0	0
Total	4	29	21	80	134

By dividing the entries in rows 1-4 in Table 2 by the corresponding row totals we get estimates of the conditional probabilities $P[T_c \in (p,q]|T_s \in (a,b]]$, and in the last rows the unconditional probabilities $P[T_c \in (p,q]]$. These probabilities, times 100%, are presented in Table 3.

Table 3. Estimated conditional probabilities

$P[T_c \in (p,q]|T_s \in (a,b]] \times 100\%$

$T_s \diagdown T_c$	$(0,1]$	$(1,6]$	$(6,12]$	$(12,\infty)$
$(0,1]$	5	21	16	58
$(1,6]$	3	22	19	56
$(6,12]$	3	20	18	59
$(12,\infty)$	0	20	20	60
$(0,\infty)$	4	21	18	57

Treatment group only:

$T_s \diagdown T_c$	$(0,1]$	$(1,6]$	$(6,12]$	$(12,\infty)$
$(0,1]$	5	21	17	57
$(1,6]$	3	23	20	54
$(6,12]$	4	14	18	64
$(12,\infty)$	0	20	20	60
$(0,\infty)$	4	21.5	18.5	56

Control group only (? = undefined):

$T_s \diagdown T_c$	$(0,1]$	$(1,6]$	$(6,12]$	$(12,\infty)$
$(0,1]$	4	22	14	60
$(1,6]$	2	18	18	62
$(6,12]$	0	50	17	33
$(12,\infty)$	?	?	?	?
$(0,\infty)$	3	21.6	15.7	59.7

Comparing the entries in rows 1-4 of Table 3 with the corresponding entries in row 5, it appears that for both groups separately and together all but one of the estimates of $P[T_c \in (p,q]|T_s \in (a,b]]$ are close to the estimates of $P[T_c \in (p,q]]$. The exception is the estimate of $P[T_c > 12|T_s \in (6,12]]$ for the control group, but this estimate is based on only two observations.

For our analysis only the probabilities $P[T_c \in (p,q]|T_s \in (a,b]]$ for $p \geq b$ are relevant. In the Appendix we present the results of tests of the null hypothesis that $P[T_c \in (p,q]|T_s \in (a,b]] = P[T_c \in (p,q]]$ for $p \geq b$. In all cases the null hypothesis is not rejected at any conventional significance level!

Therefore, it seems that the events $T_c \in (p,q]$ and $T_s \in (a,b]$ for $b \leq p$ are independent. However, if there is a treatment effect one would expect that in the case $G = 1$ these events are dependent.

These results are compatible with independence of T_c and T_s conditional on the covariates, provided that the vector X of covariates can be partitioned as

$$X = (X_s', X_c')', \text{ where } X_s \text{ and } X_c \text{ are } \textbf{independent}, \tag{1}$$

and

$$P\left[T_s \leq t|X\right] = P\left[T_s \leq t|X_s\right], \; P\left[T_c \leq t|X\right] = P\left[T_c \leq t|X_c\right]. \tag{2}$$

Therefore, at least for the control group, we will assume that the conditions (1) and (2) hold.

If there is a treatment effect, then for $t > T_s$,

$$P\left[T_c \leq t|T_s, X_c\right] \neq P\left[T_c \leq t|X_c\right], \tag{3}$$

so that for the treatment group,

$$\begin{aligned} &P\left[T_c \in (p,q], T_s \in (a,b]\right] = P\left[T_c \in (p,q]\right] P\left[T_s \in (a,b]\right] \\ &+ \int_a^b \left(P\left[T_c \in (p,q]|T_s = t_s\right] - P\left[T_c \in (p,q]\right]\right) dP\left[T_s \leq t_s\right]. \end{aligned}$$

If the latter term is small then the events $T_c \in (p,q]$ and $T_s \in (a,b]$ are approximately independent. Thus, if the dependence of $P\left[T_c \leq t|T_s, X_c\right]$ on T_s is substantially reduced after X_c is integrated out, then the inequality (3) may no longer be detectable. Therefore, a treatment effect may still be possible, but if so it will likely work via the covariates X.

6 Modeling strategy and empirical results

6.1 The Abbring-Van den Berg model

Abbring and Van den Berg (2003) consider the problem of identification of treatment effects in a bivariate mixed proportional hazard model, where one duration, S, is the timing of an intervention on another duration Y. In their Model 1a they specify the hazard functions of these duration as

$$\theta_S(t|X,V) = \lambda_S(t)\varphi_S(X)V_S \tag{4}$$

for the duration S and

$$\theta_S(t|S,X,V) = \begin{cases} \lambda_Y(t)\varphi_Y(X)V_Y & \text{if} \quad t \leq S \\ \lambda_Y(t)\varphi_Y(X)\delta\left(t|S,X\right)V_Y & \text{if} \quad t > S \end{cases} \tag{5}$$

for the duration Y, where X is a vector of covariates with support $\mathbb{X}$, $V = (V_S, V_Y)' \in (0,\infty)\times(0,\infty)$ is a vector of dependent unobserved heterogeneity variables that are independent of X, the $\lambda_i(t)$ and $\varphi_i(X)$, $i = S, Y$, are the baseline and systematic hazards, respectively, and $\delta\left(t|S,X\right)$ represents the conditional treatment effect. Implicit in this specification is the "no anticipation" condition[2]

$$\theta_S(t|s_1,X,V) = \theta_S(t|s_2,X,V) \text{ if } t \leq \min\left(s_1,s_2\right)$$

The focus in Abbring and Van den Berg (2003) is on nonparametric identification of the baseline and systematic hazards $\lambda_i(t)$ and $\varphi_i(X)$, $i = S, Y$, the treatment function $\delta(.)$ and the joint distribution $G(v)$ of V, rather than on estimation. In particular, they show that this model is nonparametrically identified if

$$\left\{(\varphi_S(x), \varphi_Y(x))'; x \in \mathbb{X}\right\} \text{ contains an open set in } \mathbb{R}^2, \tag{6}$$

and $E\left[V_S\right] < \infty,\ E\left[V_Y\right] < \infty$.

In the case of the ESEO program, S is the job search duration T_s and Y is the recidivism duration T_c. Since T_s and T_c are interval censored, and the support $\mathbb{X}$ of the covariates is countable, which violates condition (6), we cannot take unobserved heterogeneity into account. See Bierens (2007, Section 9). Besides, the effective sample size is too small for semi-nonparametric estimation of $G(v)$. See Bierens and Carvalho (2007, Section 3.4). Thus, we have to set $V_S = V_Y = 1$ in (4) and (5).

We specify the proportional hazard of job search T_s similar to (4), with $V_S = 1$, and we specify the conditional hazard of the recidivism duration T_c for the control group ($G = 0$) as well as for the treatment group ($G = 1$) in the case $t \leq T_s$ similar to (5) for the case $V_Y = 1$, $t < S$. Moreover, we specify the systematic hazards parametrically in the usual way, as the exp(.) of linear combinations of the covariates. Thus, in our notation,

$$\theta_s(t|X) = \exp\left(\beta_s' X_s\right)\lambda_s(t) \tag{7}$$

[2] Abring and Van den Berg (2003, Assumption 1).

is the conditional hazard of the job search duration T_s, with X_s a subvector of covariates relevant for job search, and

$$\theta_c(t|T_s, X, G) = \begin{cases} \exp(\beta_c' X_c)\,\lambda_c(t) & \text{if } t \leq T_s \\ \exp(\beta_c' X_c)\,\delta(t|T_s, X, G)\,\lambda_c(t) & \text{if } t > T_s \end{cases} \tag{8}$$

is the conditional hazard of the recidivism duration T_c, with X_c a subvector of covariates relevant for recidivism, where $\delta(t|T_s, X, G) = 1$ for the control group $G = 0$, and $\delta(t|T_s, X, G)$ is be determined for the treatment group $G = 1$. This specification is in accordance with our finding in the previous section that for the control group the durations T_s and T_c seem to be independent.

Given the proportional hazard structure of the model, the subvectors X_s and X_c of the covariates and the baseline hazards $\lambda_s(t)$ and $\lambda_c(t)$ will be specified in a data-driven way. Since the duration T_s and T_c are interval-censored, there is in first instance no need to specify the baseline hazards parametrically. This enables us to let the data determine how the baseline hazards look like, and what the relevant subvectors X_s and X_c are. Only the treatment effect factor $\delta(t|T_s, X, G)$ has to be specified parametrically.

6.2 Job search

Given the hazard (7) the conditional survival function of T_s is

$$S_s(t|X) = \exp\left(-\exp\left(\beta_s' X_s\right)\Lambda_s(t)\right),$$

where $\Lambda_s(t) = \int_0^t \lambda_s(\tau)\,d\tau$ is the integrated hazard. In first instance we select for X_s all available covariates $X_s = X$. Then

$$\begin{aligned} &P\left[T_s \in (a, b]|X\right] = S_s(a|X) - S_s(b|X) \\ &= \exp\left(-\exp\left(\beta_s' X\right)\Lambda_s(a)\right) - \exp\left(-\exp\left(\beta_s' X\right)\Lambda_s(b)\right). \end{aligned}$$

Since we can only estimate $\Lambda_s(t)$ for $t \in \{1, 6, 12\}$, we may without loss of generality assume that $\Lambda_s(t)$ is piecewise linear:

$$\begin{aligned} \Lambda_s(t|\alpha_s) &= \sum_{k=1}^{i-1} \alpha_k (b_k - b_{k-1}) + \alpha_i (t - b_{i-1}) \text{ for } t \in (b_{i-1}, b_i], \\ b_0 &= 0, b_1 = 1, b_2 = 6, b_3 = 12 \\ \alpha_i &> 0 \text{ for } i = 1, ..., 3,\ \alpha_s = (\alpha_1, \alpha_2, \alpha_3)' \end{aligned} \tag{9}$$

Note that $\Lambda_s(t|\alpha_s)$ is homogenous of degree one in $\alpha_s : \Lambda_s(t|c.\alpha_s) = c.\Lambda_s(t|\alpha_s)$. Therefore, we cannot allow a constant 1 in X.

In first instance we have included all available covariates in the job search model. Then we conduct a series of Wald and likelihood ratio tests to determine the subvector X_s of covariates that are relevant for the job search duration. Moreover, on the basis of the estimation results for the piecewise linear integrated baseline hazard (9) we deduct the functional form of the underlying smooth baseline hazard $\lambda_s(t)$. The details of this specification analysis can be found in the Appendix. We find that only the location dummy variables matter for job search, so that $X_s = (CHICAGO, SANDIEGO)'$. Moreover, we cannot reject the null hypothesis that the baseline hazard is of the Weibull type. Thus, the survival function now takes the form

$$\begin{aligned} S_s(t|X) &= \exp\left(-\exp\left(\beta_s' X_s\right)\alpha_{1,s} t^{\alpha_{2,s}}\right) \\ &= \exp\left(-\exp\left(\beta_s' X_s + \ln\left(\alpha_{1,s}\right)\right) t^{\alpha_{2,s}}\right). \end{aligned} \tag{10}$$

The estimation results for this model are presented in Table 4.

Table 4. Job search

Covariates	Estimates	t-val.	Parameters	Estimates	t-val.
CHICAGO	−1.228695	−7.988	$\alpha_{1,s}$	1.197083	10.880
SANDIEGO	−0.379488	−2.745	$\alpha_{2,s}$	0.884122	15.957
Log-likelihood:	−422.063		Sample size:	503	

The significant negative signs of the location dummies indicate that the job search takes longer in Chicago and San Diego than in Boston, and in Chicago longer than in San Diego. Apparently the ex-convicts in Boston receive more or better help with the job search than in the other two cities.

The results in Table 4 are only final with respect to the model specification. The coefficients involved will be re-estimated jointly with those of the recidivism model specified below.

6.3 Recidivism of the control group

In view of the results of the preliminary data analysis, we will assume that conditional on the covariates the recidivism duration T_c for the control group is independent of the job search duration T_s. Moreover, similar to the job search case, the conditional survival function of T_c for the control group will be modeled as a proportional hazard model:

$$S_c(t|X) = \exp\left(-\exp\left(\beta_c' X\right)\Lambda_c(t)\right),$$

where $\Lambda_c(t)$ is the integrated hazard. Again, we may without loss of generality assume that $\Lambda_c(t)$ is piecewise linear, as in (9).

We have followed the same specification strategy as for job search. The details can be found in the Appendix. Surprisingly, we find that none of the covariates matter for recidivism, and that the baseline hazard is constant. Thus, the distribution of T_c for the control group is exponential, so that the survival function involved takes the form

$$S_c(t|X) = \exp\left(-\alpha_c.t\right). \tag{11}$$

The maximum likelihood estimation result for α_c is:

Table 5. Recidivism $(G = 0)$

Parameter	Estimate	t-val.
α_c	0.043681	7.396
Log-likelihood:	-139.357	
Sample size:	134	

Note that this result implies that $E\left[T_c|G=0\right] = 1/\alpha_c \approx 23$ months, and that approximately,

$$\begin{array}{lcl} P\left(T_c \in (0,1]|G=0\right) & \approx & 0.04274 \\ P\left(T_c \in (1,6]|G=0\right) & \approx & 0.18781 \\ P\left(T_c \in (6,12]|G=0\right) & \approx & 0.17740 \\ P\left(T_c > 12|G=0\right) & \approx & 0.59205 \end{array}$$

6.4 Incorporating conditional treatment

For the treatment group ($G = 1$), treatment is only received if $T_c > T_s$. Therefore we will assume that if $T_c \le T_s$ the distribution of T_c is the same as for the control group:

$$P\left[T_c \le t|T_s, X, G=1\right] = 1 - \exp\left(-\alpha_c.t\right) \text{ if } t \le T_s.$$

See (11). This is the "no anticipation" condition in Abbring and Van den Berg (2003, Assumption 1). Moreover, recall from the results of the preliminary data analysis that if there is a treatment effect then it will likely work via the covariates. Therefore, let

$$P\left[T_c \leq t | T_s, X_c, G=1\right] = 1 - \exp\left(-\alpha_c . T_s\right)$$
$$\times \exp\left(-\alpha_c . \exp\left(\beta_c' X_c\right) . \left(t - T_s\right)\right) \text{ if } t > T_s,$$

where X_c is the vector of covariates involved, which now **also includes** 1 for the constant term. Thus, the conditional survival function of T_c given T_s, X_c, and G is specified as

$$S_c(t|T_s, X_c, G) = P\left[T_c > t | T_s, X_c, G\right] = I\left(t \leq T_s\right) \exp\left(-\alpha_c . t\right) \tag{12}$$
$$+ I\left(t > T_s\right) \exp\left(-\alpha_c . T_s\right) . \exp\left(-\alpha_c . \exp\left(G . \beta_c' X_c\right) . \left(t - T_s\right)\right),$$

where $I(.)$ is the indicator function.[3]

Note that the corresponding conditional hazard takes the form (8):

$$\begin{aligned} \theta_c(t|T_s, X, G) &= \frac{-\partial S_c(t|T_s, X_c, G)/\partial t}{S_c(t|T_s, X_c, G)} \\ &= \alpha_c \left(1 + I\left(t > T_s\right)\left(\exp\left(G . \beta_c' X_c\right) - 1\right)\right), \end{aligned}$$

where the systematic hazard $\exp\left(\beta_c' X_c\right)$ is now equal to 1, the baseline hazard $\lambda_c(t)$ is equal to the constant α_c, and

$$\delta\left(t | T_s, X, G\right) = 1 + I\left(t > T_s\right)\left(\exp\left(G . \beta_c' X_c\right) - 1\right).$$

Next, rewrite the survival function of T_s as

$$S_s(t|X) = \exp\left(-\exp\left(\beta_s' X_s\right) t^{\alpha_s}\right), \tag{13}$$

where now $X_s = (CHICAGO, SANDIEGO, 1)'$ and $\alpha_s = \alpha_{2,s}$. See Table 4. Then it follows from (12) and (13) that for $0 \leq a < b \leq p < q$,

$$\begin{aligned} &P\left[T_c \in (p, q], T_s \in (a, b] | X, G\right] \\ &= \int_a^b S_c(t|t_s, X_c, G) d\left(-S_s(t_s|X_s)\right) \\ &= -\int_a^b \exp\left(-\alpha_c . \left(1 - \exp\left(G . \beta_c' X_c\right)\right) t_s\right) d\left(\exp\left(-\exp\left(\beta_s' X_s\right) t_s^{\alpha_s}\right)\right) \\ &\times \left(\exp\left(-\alpha_c . \exp\left(G . \beta_c' X_c\right) . p\right) - \exp\left(-\alpha_c . \exp\left(G . \beta_c' X_c\right) . q\right)\right) \end{aligned}$$

[3] $I(true) = 1,\ I(false) = 0.$

$$= \int_{S_s(b|X_s)}^{S_s(a|X_s)} \exp\left[-\alpha_c . \exp\left(-\alpha_s^{-1}\beta_s' X_s\right)\right.$$
$$\left. \times \left(1 - \exp\left(G.\beta_c' X_c\right)\right)\left(\ln\left(1/u\right)\right)^{1/\alpha_s}\right] du$$
$$\times \left(\exp\left(-\alpha_c . \exp\left(G.\beta_c' X_c\right).p\right) - \exp\left(-\alpha_c . \exp\left(G.\beta_c' X_c\right).q\right)\right).$$

The parameters involved can now be (re-)estimated by maximum likelihood.[4]

Note that if there is a treatment effect then the effect is positive, in the sense that treatment reduces the risk of recidivism, if for $t > T_s$,

$$P[T_c > t | T_s, X_c, G = 1] > P[T_c > t | T_s, X_c, G = 0],$$

which is the case if $\beta_c' X_c < 0$.

6.5 Joint maximum likelihood results

In first instance we have chosen for X_s in (12) the vector of all available covariates, including 1 for the constant term. Again, we have conducted a series of Wald and likelihood ration test to remove insignificant covariates. See the Appendix for the details. The result is that only two covariates matter for treatment: Age and the location dummy Boston:

Table 6. Job search, recidivism and treatment effects

Job search			**Recidivism**		
Weibull parameter	Estimates	t-val.	Parameter	Estimates	t-val.
α_s	0.875049	12.320	α_c	0.041905	7.664
Covariates			Covariates		
$CHICAGO$	−1.225155	−7.440	$BOSTON$	0.425046	2.304
$SANDIEGO$	−0.324369	−2.141	AGE	−0.045503	−2.720
1	0.184070	1.868	1	1.222241	2.536
Log-Likelihood:	−847.358		Sample size:	503	

Note that the estimation results for job search are very close to the corresponding estimates in Table 4, as expected. Moreover, observe that the estimate of α_c in Table 6 is close to the estimate of α_c in Table 5.

[4] This has been done via the user-defined maximum likelihood module in *EasyReg International*. See Bierens (2006) for the latter.

6.6 Comparison with the preliminary data analysis

In the preliminary data analysis we have argued that if the dependence of $P[T_c \leq t|T_s, X_c]$ on T_s is substantially reduced after X_c is integrated out, then the inequality (3) is no longer detectable. To verify this conjecture, we have estimated $P[T_c \in (p,q]|T_s = t_s, X_c, G = 1]$ for $t_s < p$ on the basis of the results in Table 6 and then averaged these estimates over the treatment group, which yield the following results:

Table 7. Estimated $P[T_c \in (p,q]|T_s = t_s, G = 1]$

p	q	Range of t_s	Mean	Minimum	Maximum
1	6	$0 \to 1$	0.20697	0.20567	0.20829
6	12	$0 \to 1$	0.18507	0.18405	0.18610
6	12	$1 \to 6$	0.19164	0.18610	0.19753
12	∞	$0 \to 1$	0.56324	0.56194	0.56457
12	∞	$1 \to 6$	0.57202	0.56457	0.58015
12	∞	$6 \to 12$	0.59190	0.58015	0.60480

Indeed, the dependence of $P[T_c \in (p,q]|T_s = t_s, G = 1]$ on $t_s < p$ is weak, which explains why we could not find any dependence.

6.7 Treatment effects

It is straightforward to verify from (12) that

$$\begin{aligned}&P[T_c \leq t|T_c > T_s, T_s, X_c, G] \\ &= 1 - \exp\left(-\alpha_c.\exp\left(G.\beta_c' X_c\right).(t - T_s)\right) I\left(t > T_s\right)\end{aligned}$$

hence

$$E[T_c|T_c > T_s, T_s, X_c, G] = T_s + \alpha_c^{-1}\exp\left(-G.\beta_c' X_c\right)$$

and thus

$$\begin{aligned}E[T_c|T_c &> T_s, T_s, X_c, G = 1] - E[T_c|T_c > T_s, T_s, X_c, G = 0] \\ &= \alpha_c^{-1}\left(\exp\left(-\beta_c' X_c\right) - 1\right).\end{aligned} \tag{14}$$

This expression may be interpreted as (a version of) the conditional treatment effect. Thus, the treatment has a positive effect, in the sense that treatment increases the expected time between release and rearrest, if $\beta_c' X_c < 0$, regardless the job search duration.

It follows from the results in Table 6 that

$$\widehat{\beta}_c' X_c = 1.222241 + 0.425046.BOSTON - 0.045503.AGE,$$

hence in San Diego and Chicago the treatment reduces the risk of recidivism if

$$AGE > \frac{1.222241}{0.045503} \approx 27,$$

and in Boston if

$$AGE > \frac{2.180753}{0.045503} \approx 36.$$

Thus, in general, treatment only reduces the risk of recidivism for older ex-inmates, and increases the risk of recidivism for younger ex-inmates! With how much is illustrated in Figure 1

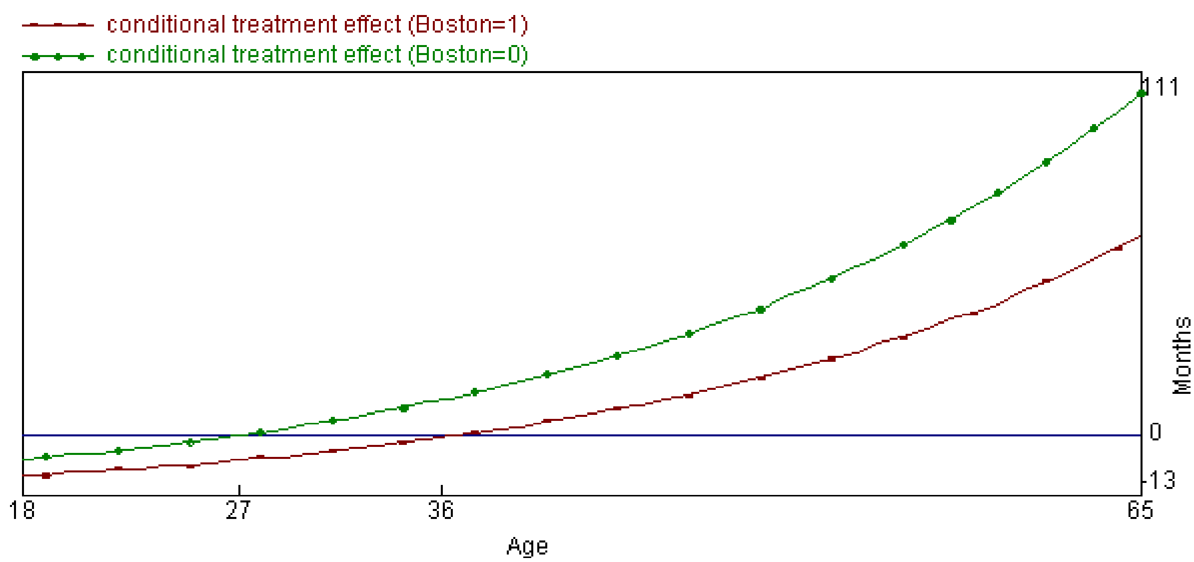

Figure 1: Conditional treatment effect on recidivism.

7 Conclusions

In contrast with previous studies we find that the ESEO program has an effect on recidivism, but this effect depends on age and location: the ESEO program reduces the risk of recidivism only for ex-inmates over the age of 27 in San Diego and Chicago, and over the age of 36 in Boston, but increases the risk of recidivism for the other ex-inmates in the treatment group. However, in view of Figure 1 is seems that the positive effect of the treatment for the older ex-convicts outweighs the negative effect for the younger ex-convicts, in terms of the expected number of months with which the rearrest will be postponed.

One of the agreements in the literature on program evaluation is that given the specificities of the groups of people who usually make use of those services, some programs that work well for a given group may not work so well for others. In other words, the effects of programs may be heterogenous. See, for example, Heckman et al. (1999). That is exactly what we find.

Recall from Tables 1 and 4 that the attrition rate and the job search duration in San Diego and Chicago are significant higher than in Boston. We hypothesized that the reason may be that the standard service in assisting the ex-inmates with the job search is much more effective in Boston than in San Diego and Chicago, and that therefore the motivation for participation in the program is higher in Boston than in the other two locations. On the other hand, in view of our results it seems at first sight that the post-placement service rendered to the ex-inmates in the treatment group in Boston is less substantial than in the other locations. However, as we will argue below, the treatment effect in Boston may be partly due to a better post-placement service!

In unemployment duration studies, age and education are usually important factors for the length of the unemployment spell. In the case under review, however, the job search duration does not depend on any individual-specific covariates. This may be due to the fact that all individuals in the sample have one dominant characteristic in common, namely being ex-convicts. Moreover, our finding that the job search duration only depends on the location may also indicate that this duration mainly measures the efforts of the program staff in the three locations in finding jobs for the ex-inmates, rather than the efforts of the ex-inmates themselves.

The available jobs for ex-inmates are likely low wage jobs, and are probably not very desirable so that employers cannot fill them with regular employees. Now suppose that a typical ex-inmate in the program is not eager to find a job, and is not actively participating in the job search, but is staying in the program for the financial aid received during the time the program staff is trying to find a job for him (or her). Moreover, suppose that the financial aid during the job search does not affect his tendency to commit crimes. Since recidivism is measured as the time between release and rearrest, it is possible that crimes are committed while the ex-inmate is in the program but for which he has not yet been caught. Once a job is found for him, the free money dries up. Since the job will likely not pay enough to be an alternative for illegal income sources, it is merely a nuisance rather than a blessing and will not last long.

However, if the ex-inmate is in the treatment group, he will for a period of 180 days be monitored and bothered by the program staff to stay in a low paying job he does not like. Also, in view of the effect of the group assignment on attrition it is conceivable that the ex-inmate had a higher expectation of the quality of the job than the one he actually gets. Anyhow, being monitored may be the answer key for the treatment puzzle. In order for the program staff to do their job, the employer has to give feedback about the performance and behavior of the ex-inmate. Now suppose that the ex-inmate shows up on the job one day in an expensive car, say a Mercedes or BMW. The employer will report that to the program staff, who will report this to the police, because the wage the ex-inmate receives will certainly not be sufficient to make car payment on this type of vehicle, so it must be stolen or financed by criminal activities. Age may fit in this story because younger people are less careful in hiding material things they are not supposed to own, or are more inclined to show off these things, than older people, or if older people have less desires for material things. Also, older ex-inmates may be more inclined to give the job a try and postpone criminal activities for a while, or are more experienced in avoiding being caught for crimes. The "Boston effect" may fit in if the ex-inmates in Boston are more closely monitored than in the other locations, so that suspicious behavior is more often detected.

Admittedly, this explanation is speculative and quite cynical. Whether it is close to the truth cannot be verified from the data. The only way to verify our hypotheses is to conduct detailed case studies in the three program locations, which is beyond our scope.

References

Abbring, J. H. and G. J. van den Berg (2003): The non-parametric identification of treatment effects in duration models, *Econometrica*, 71, 1491-1517.

Abbring, J. H., G. J. van den Berg, and J. C. van Ours (2005): The effect of unemployment insurance sanctions on the transition rate from unemployment to employment, *Economic Journal*, 115, 602-630.

Bierens, H. J. (2006): *EasyReg International*, Pennsylvania State University (http://econ.la.psu.edu/~hbierens/EASYREG.HTM).

Bierens, H. J. (2007): Semi-nonparametric interval-censored mixed proportional hazard models: Identification and consistency results, *Econometric Theory,* forthcoming (http://econ.la.psu.edu/~hbierens/SNPMPHM.PDF).

Bierens, H.J. and J. R. Carvalho (2007): Semi-nonparametric competing risks analysis of recidivism, *Journal of Applied Econometrics,* forthcoming (http://econ.la.psu.edu/~hbierens/RECIDIVISM2.PDF).

Bloom, H. S. (1984): Accounting for no-shows in experimental evaluation design, *Evaluation Review*, 8, 225 246.

Carvalho, J. R. and H.J. Bierens (2007): Conditional treatment and its effect on recidivism, *Brazilian Review of Econometrics,* forthcoming (http://econ.la.psu.edu/~hbierens/CONDTREAT.PDF).

Beck, A. J., and B. E. Shipley (1989): Recidivism of prisoners released in 1983, Special report, Bureau of Justice Statistics.

Devine, T., and N. Kiefer (1991): *Empirical Labor Economics: The Search Approach.* Oxford University Press, New York.

Eberwein, C., J. Ham, and R. Lalonde (1997): The impact of being offered and receiving classroom training on the employment histories of disadvantaged women: Evidence from experimental data, *Review of Economic Studies*, 64, 655 682.

Gendreau, P., T. Little, and C. Goggin (1996): A meta-analysis of the predictors of adult offender recidivism: What works!, *Criminology*, 34, 575 607.

Heckman, J., R. J. Lalonde, and J. A. Smith (1999): The economics and econometrics of active labor market programs, in *Handbook of Labor Economics, Vol. 3A*, Elsevier Science, Amsterdam, 1865-2097.

Heckman, J., J. Smith, and C. Taber (1998): Accounting for dropouts in evaluations of social programs, *The Review of Economics and Statistics*, 80, 1-14.

Maltz, M. D. (1984): *Recidivism*: *Quantitative Studies in Social Sciences.* Academic Press, Orlando, FL.

Milkman, R. (2001): Employment services for ex-offenders, 1981-1984: Boston, Chicago, and San Diego, Discussion Paper 8619, ICPSR.

Milkman, R., A. Timrots, A. Peyser, M. Toborg, B. G. A. Yezer, L. Carpenter, and N. Landson (1985): Employment services for ex-offenders field test, Discussion paper, The Lazar Institute.

Pissarides, C. A. (2000): *Equilibrium Unemployment Theory* (Second edition). The MIT Press, Cambridge, MA.

Schmidt, P., and A. D. Witte (1988): *Predicting Recidivism Using Survival Models*: *Research in Criminology.* Springer-Verlag, Berlin.

Timrots, A. (1985): An evaluation of employment services programs for ex-offenders, Masters thesis, University of Maryland, College Park.

Van den Berg, G. J. (1999): Empirical inference with equilibrium search models of the labor market, *The Economic Journal*, 109, 283-306.

Van den Berg, G. J. (2000): Duration models: Specification, identification, and multiple durations, in *Handbook of Econometrics, Vol. V*, North Holland, Amsterdam.

Van den Berg, G. J., B. van der Klaauw and J. C. van Ours (2004): Punitive sanctions and the transition rate from welfare to work, *Journal of Labor Economics*, 22, 211-241.

8 Appendix

8.1 Attrition

The initial Logit results for attrition are presented in Table A.1.

Table A.1. Initial Logit results for attrition

Covariates	Estimates	t-val.		Covariates	Estimates	t-val.	
AGE	0.0221627	1.08	*	$AGE1ARR \times G$	0.0730824	2.09	*
$AGE1ARR$	−0.0404668	−1.36	*	$DRUGS \times G$	−0.0553598	−0.18	*
$DRUGS$	0.1204060	0.48	*	$WHITE \times G$	0.4943870	1.55	*
$WHITE$	−0.2757420	−1.08	*	$MALE \times G$	−0.9730099	−2.16	
$MALE$	1.0069009	2.85		$CHICAGO \times G$	−1.3287316	−3.79	
$CHICAGO$	1.8851845	6.92		$SANDIEGO \times G$	−0.8245169	−2.07	
$SANDIEGO$	1.2054894	3.67		$EDUC\_L \times G$	0.0991291	0.20	*
$EDUC\_L$	0.1782564	0.48	*	$EDUC\_H \times G$	−0.0063107	−0.01	*
$EDUC\_H$	−0.3512399	−0.91	*	G	−0.7824471	−0.82	*
$AGE \times G$	−0.0110983	−0.45	*	1	−0.9178355	−1.17	
Log-likelihood:	−647.325			Sample size:	1074		

The parameters indicated by an asterix (*) are individually insignificant at the 5% significance level. The Wald test that they are also jointly zero does not reject the null hypothesis at any conventional significance level. Therefore, we have re-estimated the model without these covariates:

Table A.2. Logit results for attrition: Model 2

Covariates	Estimates	t-val.	Covariates	Estimates	t-val.
$MALE$	0.8708543	3.51	$MALE \times G$	−0.7345117	−3.41
$CHICAGO$	1.8323490	7.84	$CHICAGO \times G$	−1.3317171	−4.58
$SANDIEGO$	1.0773423	3.59	$SANDIEGO \times G$	−0.6531673	−1.88
			1	−0.8711105	−3.78
Log-likelihood:	−654.658		Sample size:	1074	

The likelihood-ratio (LR) test that the model in Table A.1 can be reduced to the model in Table A.2 has p-value 0.32867. Therefore, the null hypothesis involved cannot be rejected. The Wald test of the hypothesis that the parameters of $MALE$ and $MALE \times G$ add up to zero has p-value 0.55700, hence we may replace these two covariates with $MALE \times (1 - G)$ only, with approximately the same coefficient as $MALE$ in Table A.2. Thus, males have a higher attrition rate than females, ceteris paribus, but only if selected in the control group. The Wald test of the hypothesis that the coefficients of $CHICAGO$ and $CHICAGO \times G$ add up to zero has p-value 0.01127, hence this hypothesis should be rejected. The same applies to the hypothesis that the coefficients of $SANDIEGO$ and $SANDIEGO \times G$ add up to zero: The p-value of the Wald test involved is 0.04173.

To highlight the effect of group selection on attrition we have re-estimated the attrition model once more. The final results are presented in Table 1.

8.2 Preliminary data analysis

To test whether

$$P\left[T_c \in (1,6]\right] = P\left[T_c \in (1,6] | T_s \in (0,1]\right] \tag{15}$$

$$P\left[T_c \in (6,12]\right] = P\left[T_c \in (6,12] | T_s \in (0,1], T_s \in (1,6]\right] \tag{16}$$

$$\begin{aligned} &P\left[T_c \in (12,\infty)\right] \\ &= P\left[T_c \in (12,\infty) | T_s \in (0,1], T_s \in (1,6], T_s \in (6,12]\right] \end{aligned} \tag{17}$$

we have estimated Logit models for each of these conditional probabilities, and for each group separately:

$$P\left[T_c \in (1,6] | T_s \in (0,1]\right] = F\left(\beta_{1,0} + \beta_{1,1} I\left(T_s \in (0,1]\right) + \beta_{1,2}\right) \tag{18}$$

$$\begin{aligned} &P\left[T_c \in (6,12] | T_s \in (0,1], T_s \in (1,6]\right] \\ &= F\left(\beta_{2,0} + \beta_{2,1} I\left(T_s \in (0,1]\right) + \beta_{2,2} I\left(T_s \in (1,6]\right)\right) \end{aligned} \tag{19}$$

$$\begin{aligned} &P\left[T_c \in (12,\infty) | T_s \in (0,1], T_s \in (1,6], T_s \in (6,12]\right] \\ &= F\left(\beta_{3,0} + \beta_{3,1} I\left(T_s \in (0,1]\right) + \beta_{3,2} I\left(T_s \in (1,6]\right) + \beta_{3,3} I\left(T_s \in (6,12]\right)\right) \end{aligned} \tag{20}$$

where $F(x)$ is the logistic distribution function. However, in the case $G = 0$ we have $I\left(T_s \in (0,1]\right) + I\left(T_s \in (1,6]\right) + I\left(T_s \in (6,12]\right) = 1$, so that in the case (20) we can only estimate

$$\begin{aligned} &P\left[T_c \in (12,\infty) | T_s \in (0,1], T_s \in (1,6]\right] \\ &= F\left(\beta_{3,0} + \beta_{3,1} I\left(T_s \in (0,1]\right) + \beta_{3,2} I\left(T_s \in (1,6]\right)\right) \end{aligned} \tag{21}$$

Note that we do not need to worry about misspecification of these Logit models, because the explanatory variables involved are mutually exclusive dummy variables. Therefore, the hypothesis (15) is equivalent to the hypothesis $\beta_{1,1} = 0$ in (18), the hypothesis (16) is equivalent to the hypothesis

$\beta_{2,1} = \beta_{2,2} = 0$ in (19), and the hypothesis (17) is equivalent to the hypothesis $\beta_{3,1} = \beta_{3,2} = \beta_{3,3} = 0$ in (20) if $G = 1$, and to the hypothesis $\beta_{3,1} = \beta_{3,2} = 0$ in (21) if $G = 0$. The estimation and test results are presented in Tables A.3 and A.4

Table A.3. Logit results for (18)-(20)/(21)
Treatment group ($G = 1$)

i	$\beta_{i,0}$	$\beta_{i,1}$	$\beta_{i,2}$	$\beta_{i,3}$	Wald test
	(t-value)	(t-value)	(t-value)	(t-value)	(p-value)
1	−1.2809338	−0.0391111			0.0225
	(−7.17)	(−0.15)			(0.88076)
2	−1.5040774	−0.0606246	0.1094842		0.88076
	(−3.33)	(−0.12)	(0.22)		(0.83261)
3	0.4054651	−0.1335314	−0.2595112	0.1823216	1.19
	(0.44)	(−0.14)	(−0.28)	(0.18)	(0.75601)

Table A.4 . Logit results for (18)-(20)/(21)
Control group ($G = 0$)

i	$\beta_{i,0}$	$\beta_{i,1}$	$\beta_{i,2}$	Wald test
	(t-value)	(t-value)	(t-value)	(p-value)
1	−1.2992830	0.0216225		0.0025
	(−3.99)	(0.05)		(0.96012)
2	−1.6094379	−0.1973594	0.0930904	0.35
	(−1.47)	(−0.17)	(0.08)	(0.83801)
3	−0.6931472	1.1093076	1.1826954	1.69
	(−0.80)	(1.24)	(1.29)	(0.42906)

In all cases the null hypothesis is not rejected by the Wald test (or squared t test in the case $i = 1$) at any conventional significance level!

8.3 Job search model specification

The initial maximum likelihood results for job search are presented in Table A.5. Recall that the α's are the parameters of the integrated baseline hazard (9).

Table A.5. Job search: Model 1

Covariates	Estimates	t-val.		Covariates	Estimates	t-val.	
AGE	0.001374	0.141	*	$EDUC\_L$	−0.261975	−1.393	*
$AGE1ARR$	0.011526	0.964	*	$EDUC\_H$	−0.094193	−0.592	*
$DRUGS$	−0.060380	−0.526	*	Parameters			
$WHITE$	0.128538	1.051	*	α_1	1.139577	2.788	
$MALE$	−0.206205	−1.258	*	α_2	0.806235	2.577	
$CHICAGO$	−1.188883	−7.057		α_3	1.082720	2.186	
$SANDIEGO$	−0.353766	−2.454					
Log-likelihood:	−417.622			Sample size:	503		

The Wald test of the hypothesis that the coefficients of the covariates in Table A.5 indicated by an asterix (*) are jointly zero has p-value 0.45137. The Wald test of the null hypothesis $\alpha_1 = \alpha_2 = \alpha_3$ has p-value 0.10356, so that the null hypothesis involved cannot be rejected at the 10% significance level. Recall that the latter hypothesis implies that the baseline hazard is constant. However, for the time being we will not implement the restriction $\alpha_1 = \alpha_2 = \alpha_3$. First, we will get rid of the insignificant covariates. The results are presented in Table A.6.

Table A.6. Job search: Model 2

Covariates	Estimates	t-val.	Parameters	Estimates	t-val.
$CHICAGO$	−1.211832	−7.739	α_1	1.185860	10.711
$SANDIEGO$	−0.326150	−2.303	α_2	0.847022	6.014
			α_3	1.061613	3.719
Log-likelihood:	−420.811		Sample size:	503	

The null hypothesis $\alpha_1 = \alpha_2 = \alpha_3$ is now rejected: the Wald test involved has p-value 0.01131. On the other hand, the null hypothesis $\alpha_2 = \alpha_3$ is accepted: The Wald test involved has p-value 0.39498. The latter result indicates that the baseline hazard is non-increasing. This suggests to specify a Weibull baseline hazard. The estimation results involved are presented in Table 4.

As a double-check whether the model can be reduced from the initial model in Table A.5 to the model in Table 4 we have conducted the likelihood-ratio test: The LR test involved has p-value 0.35230, hence the hypothesis cannot be rejected at any conventional significance level. Moreover, the t-test statistic of the null hypothesis $\alpha_{2,s} = 1$ has value -2.091, which is borderline significant at the 5% level for the two-sided t-test, and significant for the left-sided t-test (the corresponding left-sided p-value is 0.01826). Since it is implausible that the "hazard" of finding a job increases with the job search duration, the left-sided result prevails.

8.4 Recidivism of the control group

The initial maximum likelihood results are presented in Table A.7:

Table A.7. Recidivism of the control group: Initial model

Covariates	Estimates	t-val.	Covariates	Estimates	t-val.
AGE	-0.057819	-2.098	$SANDIEGO$	0.069299	0.165
$AGE1ARR$	-0.061837	-1.716	$EDUC\_L$	-0.103330	-0.206
$DRUGS$	0.232762	0.733	$EDUC\_H$	0.748567	1.741
$WHITE$	-0.012282	-0.036	Parameters:		
$MALE$	-0.290917	-0.721	α_1	0.366003	0.906
$EDUC\_H$	0.748567	1.741	α_2	0.609465	0.957
$CHICAGO$	0.287844	0.725	α_3	0.493359	1.010
Log-likelihood:	-134.728		Sample size:	134	

The covariate AGE is borderline significant at the 5% level; all the other covariates are insignificant at any conventional significance level. The Wald test that all the coefficients of the covariates (including the one for AGE) are zero has p-value 0.32881, hence the null hypothesis that T_c does not depend on covariates cannot be rejected at any conventional significance level. Moreover, the Wald test of the null hypothesis $\alpha_1 = \alpha_2 = \alpha_3$ has p-value 0.80083, and therefore cannot be rejected at any conventional significance level. Recall that this hypothesis is equivalent to the hypothesis that $\Lambda_c(t) = \alpha_1 t$. Thus, the distribution of T_c for the control group is exponential, without covariates!

The LR test that the model in Table A.7 can be reduced to the exponential model in Table 5 has p-value 0.59809, hence the latter cannot be rejected at any conventional significance level. Therefore, we will adopt the exponential model (11) for the recidivism of the control group.

8.5 Joint maximum likelihood results

The initial maximum likelihood estimation results are presented in Table A.8, for recidivism only.

Table A.8. Recidivism: Model 1

Covariates	Estimates	t-val.		Covariates	Estimates	t-val.	
$DRUGS$	−0.217296	−1.027	*	$EDUC\_H$	−0.330336	−0.939	*
$CHICAGO$	−0.523601	−2.073		AGE	−0.040221	−2.200	
$SANDIEGO$	−0.461512	−1.926		$AGE1ARR$	−0.030686	−1.292	*
$WHITE$	−0.278290	−1.332	*	1	2.245037	3.226	
$MALE$	−0.050078	−0.154	*	Parameter			
$EDUC\_L$	0.576838	1.874		α_c	0.041529	7.672	
Log-Likelihood:	−842.896			Sample size:	503		

The parameters in Table A.8 indicated by an asterix (*) are jointly insignificant: The p-value of the Wald test involved is 0.42136. Therefore, in the next estimation round these covariates have been removed:

Table A.9. Recidivism: Model 2

Covariates	Estimates	t-val.	Parameter	Estimates	t-val.
$CHICAGO$	−0.301132	−1.350	α_c	0.040407	7.550
$SANDIEGO$	−0.460476	−2.030			
$EDUC\_L$	0.487441	1.639			
AGE	−0.047794	−2.870	Log-Likelihood:	−845.992	
1	1.707907	3.581	Sample size:	503	

The Wald test that only AGE matters for recidivism yields p-value 0.03464, hence this hypothesis is rejected at the 5% significance level. The Wald test that the coefficients of the two location dummy variables are equal has p-value 0.53307, which indicates that we may replace these dummy variables by the dummy variable $BOSTON = 1 - CHICAGO - SANDIEGO$. Moreover, the test of the same hypothesis, jointly with the hypothesis that the coefficient of $EDUC\_L$ is zero, has p-value 0.20804. Therefore, we have re-estimated the model without the education dummy, and with the location dummies replaced by the dummy variable $BOSTON$. See Table 6.

As a double-check we have conducted the LR test that the initial model in Table A.8 can be reduced to the model in Table 6. The p-value of the test is 0.25816.