Conditional Treatment and Its Effect on Recidivism

Herman J. Bierens
Pennsylvania State University, USA
& Tilburg University, The Netherlands

José Raimundo Carvalho [*]
Federal University of Ceara, Brazil

April, 2004

Abstract

The objective of this paper is to evaluate the effect of the 1985 "Employment Services for Ex-Offenders" (ESEO) program on recidivism. Initially, the sample has been split randomly in a control group and a treatment group. However, 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 a multiple proportional hazard model with unobserved heterogeneity for job search and recidivism time which incorporates the conditional treatment effect. We find that the program helps to reduce criminal activity, contrary to the result of the previous analysis of this data set. This finding is important for crime prevention policy.

1   Motivation

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 the conditional feature of the treatment. The timing of the treatment was completely neglected and, as will be shown later, this is a very important characteristic of the program under evaluation

2 Characterizing the Program

2.1 Antecedents

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 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 as the main obstacle to the complete success of such programs. As singled out by Mylkman, Timrots, Peyser, Toborg, Yezer, Carpenter, and Landson (1985): *"Historically, employment services programs have severed contact with the client immediately after job placement. If any follow-up occurs, it is usually limited to periodic telephone contact with the employer to determine if the client is still employed. The programs generally cease to provide support ... virtually abandoning him [client] during this crucial time in his adjustment to life outside of the institution."*

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 U.S. Department of Justice saw this as an 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. The next subsection describes the institutional details common to all three programs[1].

2.2 Institutional Framework

There are four important institutional aspects in any employment program: the eligibility rule, the assignment (between controls and treatment) scheme, provision of treatment and outcome measurement. We postpone the two first aspects until Section 3, where the details about the available data set is discussed , since we believe to be a more appropriate place. Hence, in the ESEO program[2], after being assigned to either the control or treatment

---

[1]Of course, no program was identical to the others. However, specific attributes were not relevant to deserve separate analysis.

[2]We closely follow Mylkman, Timrots, Peyser, Toborg, Yezer, Carpenter, and Landson (1985).

group[3] 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 of the provision of treatment is the job search assistance. This is the traditional job search assistance type of service, as described by Heckman, Lalonde, and Smith (1999). The job search assistance is the stage in the ESEO program that is offered equally to both controls and treatments. The difference begins upon placement. Controls were not helped after placement whereas treatments started receiving follow-up help just after the employment relation starts. The follow-up special services consisted basically of crisis intervention, counselling and, whenever necessary, reemployment assistance. These services lasted six months, and data from controls and treatments were collected at 30, 60 and 180 days after placement. For a more detailed exposition of each common service offered and received as well as those services specific of each individual program, one should consult Timrots (1985), Mylkman, Timrots, Peyser, Toborg, Yezer, Carpenter, and Landson (1985) and Milkman (2001).

3   The ESEO Program Data Set

The ESEO data set consist of 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[4] only made available 1074 usable observations: 325 in

---

[3]Unfortunately, there is no information whether clients knew their treatment status. As it will be clear later on, in the context of DT programs the state of knowledge about treatment status could help identified other parameters of interest besides the effect of treatment.

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

Boston, 489 in Chicago and 260 in San Diego. A large amount of information, sometimes very detailed, was collected from all sites. That can be broadly classified into three main categories:

- **Background variables**: demography, criminal history, employment history, educational achievement, and so on;

- **Program variables**: length of search, program participation record, reasons for drop-out, features of placement (wage, number of hours, match quality), and so on;

- **Outcome variables**: number of arrests, date of first arrest, self-reported arrests for placed people only, and so on.

A first important empirical issue is related to the characterization of the population being sampled. Unless very special assumptions are evoked, 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[5]:

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 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 on the sample induced by this selection scheme. Then, any result emerging from our econometric model must be interpreted considering those two initial issues. After this preliminary discussion, we should proceed analyzing the available sample.

---

[5]For institutional details about the ESEO program we closely followed the only 2 available published documents, i. e., Mylkman, Timrots, Peyser, Toborg, Yezer, Carpenter, and Landson (1985) and Timrots (1985) .

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 grouped, however.

The point of departure for the choice of the covariates is Schmidt and Witte (1988): age at release, time served for the sample sentence, sex, education, marital status, race, drug use, supervision status, and dummies that characterize the type of recidivism. However, we also pay close attention to the criminologic literature in recidivism, for instance Gendreau, Little, and Goggin (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). 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. For instances, Abbring, van den Berg, and van Ours (1997), Eberwein, Ham, and Lalonde (1997) and van den Berg, van der Klaauw, and van Ours (1998).

In view of those studies, a set of important covariates has been singled out. This set is composed basically of schooling, sex, age, and race. Together with the covariates related to recidivism, and the endogenous variables, the model variables are:

Endogenous variables

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

- **SEARCH**: Discrete variable indicating which interval[6] the search du-

---

[6]See Table 1.

ration belongs to. **SEARCH** = {1,2,3 or 4};

- **CRIME**: Discrete variable indicating which interval[7] the recidivism belongs to. **CRIME** = {1,2,3 or 4 };

Table 1: Duration Intervals

Number	Interval	Days
1	(0,30]	30
2	(30, 180]	150
3	(180, 360]	180
4	(360, ∞)	-

The 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 : through 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 Mylkman, Timrots, Peyser, Toborg, Yezer, Carpenter, and Landson (1985). Also, data on the first arrest after release is available. The latter is what we will use as the duration of recidivism[8]. Thus, in the sequel "(duration of) recidivism" should be interpreted as the time between release and first rearrest.

Exogenous variables

- **GROUP**: Indicator for group participation. **GROUP** = 0 means control, **GROUP** = 1 means treatment;

---

[7]See Table 1.

[8]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).

- **DRUG**: Indicator for the use of drugs during the last 5 years. DRUG = 0 means no use, DRUG = 1 otherwise;

- **RACE**: Indicator for race. RACE = 0 means white, RACE = 1 means non-white;

- **SEX**: Indicator for sex. SEX = 0 means female, SEX = 1 means male;

- **EDUC**: Discrete variable describing educational attainment. EDUC = 0 if individual has from 2 to 8 years of schooling, EDUC = 1 if he/she has from 9 to 12 years or GED, and EDUC = 2 if he/she has more than 12 years;

- **AGE**: Age of ex-convict, in years;

- **SANDIEGO**: Indicator for city. SANDIEGO = 1 means San Diego, SANDIEGO = 0 means either Chicago or Boston;

- **CHICAGO**: Indicator for city. CHICAGO = 1 means Chicago, CHICAGO = 0 means either San Diego or Boston;

- **AGEFIRST**: Age at first arrest, in years.

A summary of descriptive statistics of the covariates is given in Table 2.

Table 2: Descriptive Statistics

Variable	Observations	Min	Max	Mean	Stdev
ATTRITION	1074	0	1	0.5316	0.4992
GROUP	1074	0	1	0.5744	0.4946
DRUG	1074	0	1	0.3752	0.4844
RACE	1074	0	1	0.3352	0.4722
SEX	1074	0	1	0.8929	0.3093
EDUC	1074	0	2	1.0242	0.4529
AGE	1074	16	59	27.4497	6.5739
SANDIEGO	1074	0	1	0.2420	0.4285
CHICAGO	1074	0	1	0.4553	0.4982
MOSUNEMP	1074	0	97	9.3791	19.5649
AGEFIRST	1074	6	44	16.4543	4.5406

4 An Econometric Model of the ESEO Program

4.1 Identification of treatment effect

To the best of our knowledge, Abbring and den Berg (2000a) is one of the first attempts to model treatment effects in a context of duration analysis that rigourously discusses nonparametric identification. However, the type of treatment effect identified by these authors is not the same treatment effect used in the literature on econometric program evaluation, because Abbring and den Berg (2000a) do not consider the presence of a control group as it is traditionally present in evaluation studies. Nevertheless, the treatment effect in our model is identified, but we have established this empirically rather than theoretically.

4.2 The Model

4.2.1 Absense of treatment

The latent dependent variables in or model are T_s, the job search time since release from prison, and T_c, the time of the first arrest after release from prison. Let $V \in \mathbb{R}_+$ be a random variable representing unobserved heterogeneity. In the absence of treatment the model could be specified according the approach advocated by van den Berg (2000): conditional on the unobserved heterogeneity V and the exogenous variables in a vector X, the durations T_s and are T_c independent. Adopting a proportional representation for the hazard functions,

$$\theta_s(t|X,V) = \lambda_s(t) \cdot \phi_s(X) \cdot V. \tag{1}$$

$$\theta_c(t|X,V) = \lambda_c(t) \cdot \phi_c(X) \cdot V. \tag{2}$$

the conditional survival functions, given X and V, for each of the durations T_s, T_c are

$$S_s(t|X,V) = P(T_s \geqslant t|X,V,W) = \exp\left(-V \cdot \phi_s(X) \cdot \int_0^t \lambda_s(\tau)d\tau\right) \quad (3)$$

$$S_c(t|X,V) = P(T_c \geqslant t|X,V,W) = \exp\left(-V \cdot \phi_c(X) \cdot \int_0^t \lambda_c(\tau)d\tau\right) \quad (4)$$

Hence, the joint conditional survival function conditional on X and V is:

$$S(t_s,t_c|X,V) = P\left[T_s \geq t_s, T_c \geq t_c|X,V\right] \quad (5)$$

$$= \exp\left(-V \cdot \left(\phi_s(X) \cdot \int_0^{t_s} \lambda_s(\tau)d\tau + \phi_c(X) \cdot \int_0^{t_c} \lambda_c(\tau)d\tau\right)\right)$$

Finally, in order tighten the durations T_s, T_c together and make them dependent conditional on X only, the random variable V has to be integrated out. Given a specification $G(v)$ of the distribution function of V, the joint survival function conditional on X alone is:

$$S(t_s,t_c|X) = \mathcal{L}\left(\phi_s(X) \cdot \int_0^{t_s} \lambda_s(\tau)d\tau + \phi_c(X) \cdot \int_0^{t_c} \lambda_c(\tau)d\tau\right),$$

where $\mathcal{L}(.)$ is the Laplace transform of G:

$$\mathcal{L}(s) = \int_0^\infty \exp\left(-v.s\right) dG(v), \ \ s \geq 0.$$

4.2.2 Incorporating treatment

The key issue now is how to incorporate treatment in this framework. Let the dummy variable W represent group participation: $W = 1$ if the individual is selected in the treatment group, and $W = 0$ if selected in the control group. Then treatment is received if

1. The individual is selected in the treatment group: $W = 1$.

2. The job search has ended before the first arrest: $T_s < T_c$.

The problem is now that due to the latter condition it is impossible to build in the effect of treatment directly in the joint survival function (5) without sacrificing the conditional independence of T_s and T_c given X and

10

V. However, note that without assuming conditional independence we can still factorize out the joint density of T_s and T_c conditional on X, V, and W, as a product of conditional densities, say:

$$
\begin{aligned}
& f\left(t_s, t_c | X, V, W\right) \\
= & \ f_c\left(t_c | T_s = t_s, X, V, W\right) . f_s\left(t_s | X, V, W\right).
\end{aligned}
$$

Consequently, the corresponding joint survival function can be written as

$$
\begin{aligned}
& S(t_s, t_c | X, V, W) \\
= & \ P\left[T_c \geq t_c, T_s \geq t_s X, V, W\right] \\
= & \ \int_{t_s}^{\infty} \int_{t_c}^{\infty} f_c\left(t_c | T_s = t_s, X, V, W\right) dt_c \ f_s\left(t_s | X, V, W\right) dt_s
\end{aligned}
$$

Therefore, in modeling the joint survival function of T_s, T_c conditional on X, V, and W we can still use a similar setup as before, as follows.

First, model the conditional hazard function of T_c conditional on $T_s = t_s, X, V, W$ as

$$
\begin{aligned}
& \theta_c(t | t_s, X, V, W) \\
= & \ [(1 - W)\phi_c(X) + W.(1 - I(t > t_s))\phi_c(X) + W.I(t > t_s)\phi_c^*(X)] \\
& \times \lambda_c(t_c) \cdot V. \\
= & \ [\phi_c(X) + W.I(t > t_s)\left(\phi_c^*(X) - \phi_c(X)\right)] \times \lambda_c(t_c) \cdot V.
\end{aligned}
$$

where $I(.)$ is the indicator function. If $W = 0$ this specification corresponds to the previous one in (2), but for $W = 1$ the effect of the treatment on recidivism is now incorporated:

$$
\begin{aligned}
\theta_c(t_c | t_s, t_a, X, V, W & = 1) = \phi_c^*(X) \cdot \lambda_c(t_c) \cdot V \ if \ t_s < t_c \\
\theta_c(t_c | t_s, t_a, X, V, W & = 1) = \phi_c(X) \cdot \lambda_c(t_c) \cdot V \ if \ not,
\end{aligned}
$$

where $\phi_c(X)$ is the same as in (2), and $\phi_c^*(X)$ is the systematic hazard during treatment. The corresponding conditional survival function of T_c is now

$$
\begin{aligned}
S_c(t_c | t_s, X, V, W) & = P(T_c \geqslant t_c | T_s = t_s, X, V, W) \\
& = \exp\left(-V \cdot \Lambda_c(t_c | t_s, X, W)\right)
\end{aligned}
$$

11

where

$$\Lambda_c(t_c|t_s, X, W) \tag{6}$$
$$= \phi_c(X) \int_0^{t_c} \lambda_c(\tau)d\tau + W.\left(\phi_c^*(X) - \phi_c(X)\right) \int_0^{t_c} I(\tau > t_s)\lambda_c(\tau)d\tau$$
$$= \phi_c(X) \int_0^{t_c} \lambda_c(\tau)d\tau + W.\left(\phi_c^*(X) - \phi_c(X)\right) I(t_c > t_s) \int_{t_s}^{t_c} \lambda_c(\tau)d\tau$$

is the corresponding integrated hazard.

The conditional survival function of T_s is the same as before:

$$S_s(t_s|X, V, W) = P(T_s \geqslant t_s|X, V, W)$$
$$= P(T_s \geqslant t_s|X, V) = \exp\left(-V \cdot \Lambda_s(t_s|X)\right)$$

where

$$\Lambda_s(t_s|X) = \phi_s(X) \int_0^{t_s} \lambda_s(\tau)d\tau$$

is the integrated hazard. Thus, the joint survival function of T_s, T_c conditional on X, V, and W is:

$$S(t_s, t_c|X, V, W) \tag{7}$$
$$= P\left[T_c \geq t_c, T_s \geq t_s|X, V, W\right]$$
$$= \int_{t_s}^{\infty} S_c(t_c|\tau, X, V, W) f_s\left(\tau|X, V, W\right) d\tau$$
$$= \int_{t_s}^{\infty} \exp\left[-V \cdot \Lambda_c(t_c|\tau, X, W)\right] V \exp\left(-V \cdot \Lambda_s(\tau|X)\right) \phi_s(X)\lambda_s(\tau)d\tau$$
$$= V.\phi_s(X) \int_{t_s}^{\infty} \exp\left[-V \cdot \left(\Lambda_c(t_c|\tau, X, W) + \Lambda_s(\tau|X)\right)\right] \lambda_s(\tau)d\tau$$

where the last two equalities follows from

$$f_s\left(t|X, V, W\right) = -\frac{\partial}{\partial t} S_s(t|X, V, W) \tag{8}$$
$$= -\frac{\partial}{\partial t} \exp\left(-V \cdot \Lambda_s(t|X)\right)$$
$$= V \exp\left(-V \cdot \Lambda_s(t|X)\right) \phi_s(X) \cdot \lambda_s(t)$$

12

4.2.3 Baseline hazards

The baseline hazards $\lambda_s(t)$ and $\lambda_c(t)$ are assumed to have a Weibull specification:

$$
\begin{aligned}
\lambda_s(t) &= \lambda_s t^{\lambda_s-1}, \ \lambda_s > 0, \\
\lambda_c(t) &= \lambda_c t^{\lambda_c-1}, \ \lambda_c > 0.
\end{aligned} \tag{9}
$$

For search or unemployment durations, the Weibull hazards is flexible enough to capture any pattern of monotonic dependence typical of labor markets. See for example van den Berg, Lindeboom, and Ridder (1994). Regarding criminal behavior, a "parabola" type hazards might be more appropriate. Generally, after release, the ex-criminal has a period of low criminal activity followed by a high criminal activity period. However, as long as the abscissa of the point of maximum of that parabolic hazards is close enough to the origin, the Weibull hazards is still a reasonable approximation. Then the integrated conditional hazards become

$$
\begin{aligned}
&\Lambda_c(t_c|t_s, X, W = 1) \\
&= \ \phi_c(X)t_c^{\lambda_c} + (\phi_c^*(X) - \phi_c(X))\, I(t_c > t_s)\left(t_c^{\lambda_c} - t_s^{\lambda_c}\right)
\end{aligned} \tag{10}
$$

and

$$
\Lambda_s(t|X) = \phi_s(X) \cdot t^{\lambda_s}. \tag{11}
$$

Hence, the joint conditional survival function for the treatment group takes the form:

$$
\begin{aligned}
&S(t_s, t_c|X, V, W = 1) \\
&= \ I(t_c > t_s)V.\phi_s(X)\exp\left(-V \cdot \phi_c(X)t_c^{\lambda_c}\right) \\
&\quad \times \int_{t_s}^{t_c} \exp\left[-V \cdot (\phi_c^*(X) - \phi_c(X))\left(t_c^{\lambda_c} - \tau^{\lambda_c}\right)\right] \\
&\quad \times \exp\left[-V \cdot \left(\phi_s(X) \cdot \tau^{\lambda_s}\right)\right] \lambda_s \tau^{\lambda_s-1} d\tau \\
&\quad + I(t_c > t_s)\exp\left[-V \cdot \left(\phi_c(X)t_c^{\lambda_c} + \phi_s(X) \cdot t_c^{\lambda_s}\right)\right] \\
&\quad + I(t_c \le t_s)\exp\left[-V \cdot \left(\phi_c(X)t_c^{\lambda_c} + \phi_s(X) \cdot t_s^{\lambda_s}\right)\right]
\end{aligned} \tag{12}
$$

(see Appendix A for the details of the derivations involved), whereas for control group:

$$S(t_s, t_c | X, V, W = 0) \tag{13}$$
$$= \exp\left[-V \cdot \left(\phi_c(X) \cdot t_c^{\lambda_c} + \phi_s(X) \cdot t_s^{\lambda_s}\right)\right]$$

4.2.4 Systematic hazards

For the systematic hazards $\phi_s(X)$ and $\phi_c(X)$ we adopt the usual exponential specification:

$$\begin{aligned} \phi_s(X) &= \exp(\beta_s' X), \\ \phi_c(X) &= \exp(\beta_c' X), \end{aligned} \tag{14}$$

where X contains 1 for the constant term. As to the specific hazard upon treatment, we assume that

$$\phi_c^*(X) = \delta\phi_c(X) = \delta \exp(\beta_c' X), \ \delta > 0.$$

In the sequel, however, we will continue to use the notations $\phi_s(X)$, $\phi_c(X)$ and $\phi_c^*(X)$.

The parameter δ is the key parameter on our model, as it measures the effect of the ESEO program on the recidivism behavior of its participants. The parameter δ either inflates or deflates the systematic hazards function of recidivism upon placement. Its interpretation is:

- If $\delta > 1$, the program has a negative impact on recidivism, as it inflates the hazards for recidivism, therefore shortening the time between release and first arrest;

- If $\delta = 1$, the program has no effect;

- If $\delta < 1$, the program has a positive impact on recidivism, as it deflates the hazards for recidivism, therefore lengthening the time between release and first arrest.

4.2.5 Unobserved heterogeneity

The traditional[9] choice of the distribution of the heterogeneity variable V is the Gamma distribution, because its Laplace transform has a closed form expression: If $V \sim Gamma(\alpha, \omega)$ then the Laplace transform of V is:

$$\mathcal{L}(s) = E\left[\exp(-s.V)\right] = (1 + s \cdot \omega)^{-\alpha}, \tag{15}$$

with derivative

$$\mathcal{L}'(s) = -E\left[V \exp(-s.V)\right] = -\alpha\omega \left(1 + s \cdot \omega\right)^{-\alpha-1}. \tag{16}$$

Adopting the specification it follows from (12) through (16) that:

$$\begin{aligned}
S(t_s, t_c | X, W = 1) & \tag{17}\\
= \quad & I(t_c > t_s)\alpha\omega\phi_s(X)\\
& \times \int_{t_s}^{t_c} \left[1 + \omega \left(\phi_c(X)t_c^{\lambda_c} + (\phi_c^*(X) - \phi_c(X)) \left(t_c^{\lambda_c} - \tau^{\lambda_c}\right) + \phi_s(X) \cdot \tau^{\lambda_s}\right)\right]^{-(\alpha+1)}\\
& \times \lambda_s \tau^{\lambda_s - 1} d\tau\\
& + I(t_c > t_s) \left[1 + \omega \left(\phi_c(X)t_c^{\lambda_c} + \phi_s(X) \cdot t_c^{\lambda_s}\right)\right]^{-\alpha}\\
& + I(t_c \leq t_s) \left[1 + \omega \left(\phi_c(X)t_c^{\lambda_c} + \phi_s(X) \cdot t_s^{\lambda_s}\right)\right]^{-\alpha}
\end{aligned}$$

and

$$S(t_s, t_c | X, W = 0) = \left[1 + \omega \left(\phi_c(X) \cdot t_c^{\lambda_c} + \phi_s(X) \cdot t_s^{\lambda_s}\right)\right]^{-\alpha}. \tag{18}$$

Note that ω cannot be identified. To see this, substitute (14) in 18:

$$\begin{aligned}
& S(t_s, t_c | X, W = 0) \\
= \quad & \left[1 + \exp(\ln(\omega) + \beta_c' X) \cdot t_c^{\lambda_c} + \exp(\ln(\omega) + \beta_s' X) \cdot t_s^{\lambda_s}\right]^{-\alpha}.
\end{aligned}$$

Since $X'\beta_1$ and $X'\beta_2$ contain constant terms, $\ln(\omega)$ can be absorbed in the constants. **Consequently, we will set $\omega = 1$.**

---

[9]See for example Lancaster (1990).

4.2.6 Attrition

There are two types of attrition in our sample, namely "no show" if an individual does not participate at all in the job search stage of the program, and "quitting" of an individual during the job search stage. As to attrition, we decided to take a very pragmatic approach. Instead of modelling these two types of attrition jointly with job search and recidivism, we assume that the survival functions (17) and (18) apply **conditionally** on the absence of attrition, where attrition now includes "no show" and "quitting".

If an individual quits after finding a job, and this individual is in the treatment group, we will assume that the treatment effect is the same as for an individual who completes the treatment.

Let $A = 1$ indicate attrition, and $A = 0$ absence of attrition. We will specify the probability of attrition as a Logit model:

$$P[A = 1|X, W = w] = \frac{1}{1 + \exp\left(-\gamma_w' X\right)}, \quad w = 0, 1, \tag{19}$$

The parameters γ_w may be different for $w = 0, 1$.

4.2.7 Censoring

The actual durations T_s and T_c are not directly observed, but are only known to belong to particular intervals, i.e., T_s and T_c are known to belong to one of the following four intervals appearing in Table 1.

There are 12 combinations where T_s and T_c are in different intervals: $T_s \in [a_i, b_i), T_c \in [c_i, d_i)$, say, where either $b_i \leq c_i$ or $d_i \leq a_i$. The remaining four cases, $T_s \in [a_i, b_i), T_c \in [a_i, b_i)$, will be treated as "other", because there are relatively few observations for which the latter applies, and secondly, the computation of $P(T_s \in [a_i, b_i), T_c \in [a_i, b_i))$ is more complicated than in the non-overlapping cases.

Probabilities of the type $P(T_s \in [a, b), T_c \in [c, d))$ can easily be computed on the basis of the joint survival functions:

$$\begin{aligned} P(T_s &\in [a, b), T_c \in [c, d)|X, W) \\ &= S(a, b|X, W) - S(b, c|X, W) \\ &\quad - S(a, d|X, W) + S(b, d|X, W). \end{aligned} \tag{20}$$

4.3 The likelihood function

Let $I_i = [a_i, b_i) \times [c_i, d_i)$, $i = 1, ..., k$, be **disjoint** intervals in $\mathbb{R}^2_+$. For each individual j, assign a dummy variable $D_{i,j}$ such that $D_{i,j} = 1$ if $(T_{c,j}, T_{s,j}) \in I_i$, and let $D_{0,j} = 1 - \sum_{i=1}^n D_{i,j}$. Then for $i = 1, ..., k$,

$$
\begin{aligned}
P[D_{i,j} &= 1|X_j, W_j] = P\left[(T_{c,j}, T_{s,j}) \in I_i | X_j, W_j\right] \\
&= S(a_i, b_i|X_j, W_j) - S(b_i, c_i|X_j, W_j) \\
&\quad - S(a_i, d_i|X_j, W_j) + S(b_i, d_i|X_j, W_j) \\
&= p_{i,j}(\theta),
\end{aligned}
$$

say, where

$$
\theta = (\beta'_s, \lambda_s, \beta'_c, \lambda_c, \delta, \alpha)',
$$

with $W_j = 0$ if individual j belongs to the control group, and $W_j = 1$ to the treatment group. Moreover, the probability of an individual belonging to the category "other" is:

$$
P[D_{0,j} = 1|X_j, W_j] = 1 - \sum_{i=1}^k p_{i,j}(\beta) = p_{0,j}(\beta),
$$

say. Next, let $A_j = 1$ if individual j does not show up, or quit, with probability

$$
P[A_j = 1|X_j, W_j = i] = q_j(\gamma_i), \ i = 0, 1
$$

say. See (19). Moreover, recall that we have assumed that

$$
P(T_{c,j} > a, T_{s,j} > b|X_j, W_j, A_j = 0) = S(a, b|X_j, W_j)
$$

Then the log-likelihood takes the form:

$$\log \mathbf{L}(\theta, \gamma_0, \gamma_1)$$

$$= \sum_{j=1}^{n} A_j \left((1 - W_j) \ln q_j(\gamma_0) + W_j \ln q_j(\gamma_1)\right)$$

$$+ \sum_{j=1}^{n} (1 - A_j) \left[\sum_{i=0}^{k} D_{i,j} \ln p_{i,j}(\theta) + (1 - W_j) \ln \left(1 - q_j(\gamma_0)\right) + W_j \ln \left(1 - q_j(\gamma_1)\right) \right]$$

$$= \sum_{j=1}^{n} A_j \left((1 - W_j) \ln q_j(\gamma_0) + W_j \ln q_j(\gamma_1)\right)$$

$$+ \sum_{j=1}^{n} (1 - A_j) \left[(1 - W_j) \ln \left(1 - q_j(\gamma_0)\right) + W_j \ln \left(1 - q_j(\gamma_1)\right)\right]$$

$$+ \sum_{j=1}^{n} (1 - A_j) \sum_{i=0}^{k} D_{i,j} \ln p_{i,j}(\theta)$$

$$= \log \mathbf{L}_0(\gamma_0) + \log \mathbf{L}_1(\gamma_1) + \log \mathbf{L}_2(\theta),$$

say, where n is the sample size.

4.4 Estimation and Inference

All econometric work (data manipulation, estimation and inference) was conducted by means of the econometric package **EasyReg International**[10].

4.4.1 Attrition

Since we are going to estimate both vector of regressors γ_0 and γ_1 simultaneously, we adjust the set of regressors X by way of the group dummy W. Hence, the actually estimated logit specification is:

$$P[N = 1 | X, W] = \frac{1}{1 + \exp\left(-X'\gamma_0 - W \cdot X'(\gamma_1 - \gamma_0)\right)}. \tag{21}$$

The results for the logit regression appear in Table 3. Note that given the logit specification, the second set of estimated parameters corresponds to

---

[10]This freeware package was developed by the first author and can be downloaded from: http://econ.la.psu.edu/~hbierens/EASYREG.HTM.

the difference between the estimated parameters of treatments and controls. Thus, the first set of estimates are the estimates of the components of γ_0, and the second set corresponds to $\gamma_1 - \gamma_0$. The majority of the estimated regressors are not significant at the 10% level[11].

Table 3: Logit Parameters Estimate:

Group	Parameters	ML estimate	t-values
controls	age	0.022	1.11
	first arrest	-0.040	-1.36
	drug	0.120	0.48
	race	-0.275	-1.08
	sex	1.006	2.85
	education	-0.176	-1.11
	chicago	1.884	6.99
	sandiego	1.205	3.74
	intercept	-0.741	-0.95
treatments - controls	age	0.011	0.87
	first arrest	0.032	1.76
	drug	0.062	0.34
	race	0.222	1.18
	sex	0.036	0.13
	education	-0.194	-1.64
	chicago	0.548	2.50
	sandiego	0.373	1.67
	intercept	-0.761	-0.81
	Log-Likelihood		-647.686
	n		1074

For controls, only variables sex, chicago and sandiego are significant. A man has a higher probability of attrition than a woman. Belonging to the program located in Chicago, as well as in San Diego, raises the probability of attrition. Most of the components of $\gamma_1 - \gamma_0$ are insignificant, except the dummy Chicago: for ex-inmates having served their last term in a Chicago

---

[11]For the sake of convenience, from now on, any reference made about parameter significance implicitly assumes a level of 10%.

prison, the probability of attrition is higher for the treatment group than for the control group.

4.4.2 Job search, recidivism, and treatment

Results appear in Table 4. In order to interpret the results, note that if a coefficient is positive and the corresponding X variable increases, then the whole hazards function will be inflated, hence the integrated hazard will be reduced, and so will be the survival probability. Thus, failure will occur earlier. In the case of search duration, this implies that the average time of search (unemployment) will be lower the higher the value of the X variable is. For the crime (recidivism) duration this implies that the expected time between release and rearrest will be reduced.

Some parameter estimation results for search duration appear to contradict well established facts in the literature of empirical search models. However, given the specific nature of our data (ex-criminals), there are some reasonable explanations for that. The demand side of the job market appears to be driven much more by the possibility that the future worker could commit a crime after being hired than by pure efficiency considerations. Also, the job market for ex-criminals is characterized by being of bad quality and by offering low wages. Such empirical evidence concerning job search for ex-inmates looks promising as a topic for future development.

The estimated parameter for age is negative and therefore, as expected, the job search time is higher the older the ex-inmate is.

Males appears to have search time greater than the search time of women. This is the first result that contradicts empirical findings in search models. Indeed, the sex effect is significant. However, the male/female ratio of inmates is much higher than the 50% ratio out of prisons, so that males may present a higher potential threat of comitting a crime while employed. The positive coefficient of education means that more educated ex-inmates will find jobs faster than less educated, which is in accordance with the empirical search literature. See, e.g., Devine and Kiefer (1991). The significantly negative coefficients of the dummies for Chicago and San Diego, indicate that job search time in Chicago and San Diego[12] larger than in Boston, ceteris paribus. Finally, the parameter of the baseline hazards presents a rather surprising result. As shown in Lancaster (1990), a value of $\lambda_s = 1.582$ means

---

[12]Of course, it is not possible to distinguished between local labor market conditions and program's features.

Table 4: Parameters Estimate:

Duration	Parameters	ML estimate	t-values
search	age	-0.075	-3.423
	first arrest	0.007	0.306
	drug	0.090	0.308
	race	0.167	0.583
	sex	-1.147	-2.302
	education	0.348	1.968
	chicago	-2.021	-5.540
	sandiego	-1.071	-3.021
	intercept	-2.600	-2.529
	λ_s	1.582	9.862
crime	age	-0.117	-5.320
	first arrest	-0.031	-1.319
	drug	-0.500	-1.773
	race	0.523	1.895
	sex	-1.670	-4.078
	education	0.200	1.100
	chicago	-0.361	-1.130
	sandiego	-0.026	-0.086
	intercept	-2.986	-3.147
	λ_c	1.676	10.179
heterogeneity	α	1.213	3.075
effect of treatment	δ	0.631	3.947
	Log-Likelihood		-6425.281
	n		1074

that the search time presents positive dependence, or in other words, the longer an individual keep searching the higher the probability of finding a job. This is exactly the opposite of a lot of evidence found in studies of search in the labor market. For instance, see Devine and Kiefer (1991). A plausible explanation for the high value of λ_s (which is significantly greater than 1) is that the ex-inmate will experience pretty soon after release how difficult is to find a reasonable job, given the strong stigma prevalent in the society. Hence, sooner or later he/she will have no option but to accept any

21

(low-paying) job that becomes available.

The impact of age on the expected recidivism time is significantly negative. This is in accordance with other studies in recidivism such as Schmidt and Witte (1988). Hence, *ceteris paribus*, an older ex-inmate will postpone his/her next crime. The estimated parameter of the dummy variable race (1 = black) is positive, but only borderline significant. Hence, non-whites seem to recidivate earlier than whites, which is in accordance with other empirical studies on recidivism, such as Schmidt and Witte (1988). The strongly significant negative value of the coeffient of sex appears to contradict the literature on criminal recidivism: females will commit a crime earlier than males. However, the sample consists only for about 11% of females, so that a very few bad ones among them may cause this effect. The city dummies do not have a significant effect. Also for recidivism the parameter λ_c is significantly greater than 1, which implies that the longer an ex-inmate is out without committing a crime the higher the probability of committing a crime in the future.

The parameter α is a nuisance parameter with no particular interesting interpretation other than that it is the expected value of the unobserved heterogeneity variable V.

The parameter δ is the key parameter on our econometric model. The estimated value is significantly less than 1.Hence, the program is **effective** as it increases the time between release and rearrest. This result stands in contrast with the original study of the ESEO program, as shown in Mylkman, Timrots, Peyser, Toborg, Yezer, Carpenter, and Landson (1985).

5   Conclusion

By modelling the ESEO program as a mixed multivariate proportional hazards model[13], where treatment is conditional on placement, we have merged two important fields of modern econometrics: survival analysis and econometric evaluation of programs. As far as we know, our paper is the first one to build this type of model and estimate it. The following paragraphs conclude by discussing the main achievements of the present paper, as well as by offering some possible ideas for future research.

First, our contribution has to do with the available data set. Even though this data set has been used before, it was restricted to the community of

---

[13]See, Abbring and den Berg (2000b) for this nomenclature.

sociologists and criminologists. Despite the fact that search models have been estimated since the early 80's, search by ex-inmates who participate in a program of reemployment is a novelty for the econometric audience. The estimated parameters appearing in table 4, and the discussions that followed it show that some regressors have very different effects when compared to the traditional search model. Nonetheless, our available data presents some limitations. The main limitation of our data set is that it is grouped and this definitely imposes constraint on what can be identified from the model and makes our results less convincing. A good standard to be followed by criminologists and sociologist would be the methodology used by the agencies that collect unemployment data in the USA. Better data help a lot, specially in econometric evaluation of programs, as shown by Heckman, Lalonde, and Smith (1999).

Second, we have shown some evidence of how the process of search for jobs could be heavily influenced by the demand side of the market. More specifically, it would not be surprisingly that information asymmetries play a crucial role in this specific labor market. It is very likely that all prospective employers know that each of application comes from an ex-inmate, however knowledge of the past criminal history of each ex-convicts does not need to follow. Indeed, legislation regarding disclosure of criminal past records varies a lot within the USA. Hence, an interesting topic for future research would be estimation of models that explicitly consider the information asymmetries existent in this market. We think this should be a nice starting point to address the actual debate about disclosure of criminal records and to evaluate its policy implications.

Third, the blending of survival analysis and econometric program evaluation represents our key contribution. We have set a model where the timing of treatment is explicitly considered. This stands in contrast with any other past study of econometric evaluation of programs. In fact, we are able to build an estimable model and estimate it. The estimated parameters clearly show that the timing of treatment is an important feature of social programs well neglected in the past. Nonetheless these initial accomplishments, there is still important topics for future development. Although the parameter δ serves as a general measure of program effectiveness, it is a crude measure, indeed. One of the agreements on the literature of program evaluation is that given the specificities of the groups of people who usually make use of those services, some programs that work very well for a given group could work badly for others. In other words, the effects of programs are heterogenous

and this should be accounted for.

This lead us to suggest an urgent new avenue to explore: build models that assume impact heterogeneity. From the perspective of our model this means to specify the following:

$$\delta(X) = J(X) \qquad \text{where } J(X) \geq 0 \quad \text{for all } X \in \mathbb{R}^n. \tag{22}$$

Thus, the effect of the program is conditional on a set of regressors representing individual-specific, program-specific and local variables. Undoubtedly, this would give us a much more accurate picture of the program. As a matter of fact, an easy choice would be $\delta(X) = \exp[X'\beta]$. However, identification of the model becomes a problem!

The *Handbook of Labor Economics* chapter on econometric evaluation of active labor market programs shows how well-developed and active remains this topic of research. Also, the *Handbook of Econometrics* chapter on duration analysis reaches similar conclusion. Such intersection unequivocally opens an exciting new area of research where the time dimension of programs are studied with further detail. This new approach will definitely result on the advancement of our understanding of how social programs work.

A Joint Survival Function

It follows from equations (7), (10) and (11) that:

$$
\begin{aligned}
& S(t_s, t_c | X, V, W = 1) \qquad\qquad (23) \\
=\ & V.\phi_s(X) \exp\left(-V \cdot \phi_c(X) t_c^{\lambda_c}\right) \\
& \times \int_{t_s}^{\infty} \exp\left[-V \cdot (\phi_c^*(X) - \phi_c(X))\, I(t_c > \tau)\left(t_c^{\lambda_c} - \tau^{\lambda_c}\right)\right] \\
& \times \exp\left[-V \cdot \left(\phi_s(X) \cdot \tau^{\lambda_s}\right)\right] \lambda_s \tau^{\lambda_s - 1} d\tau \\
=\ & I(t_c > t_s) V.\phi_s(X) \exp\left(-V \cdot \phi_c(X) t_c^{\lambda_c}\right) \\
& \times \int_{t_s}^{\infty} \exp\left[-V \cdot (\phi_c^*(X) - \phi_c(X))\, I(t_c > \tau)\left(t_c^{\lambda_c} - \tau^{\lambda_c}\right)\right] \\
& \times \exp\left[-V \cdot \left(\phi_s(X) \cdot \tau^{\lambda_s}\right)\right] \lambda_s \tau^{\lambda_s - 1} d\tau \\
& + I(t_c \leq t_s) V.\phi_s(X) \exp\left(-V \cdot \phi_c(X) t_c^{\lambda_c}\right) \\
& \times \int_{t_s}^{\infty} \exp\left[-V \cdot \left(\phi_s(X) \cdot \tau^{\lambda_s}\right)\right] \lambda_s \tau^{\lambda_s - 1} d\tau \\
=\ & I(t_c > t_s) V.\phi_s(X) \exp\left(-V \cdot \phi_c(X) t_c^{\lambda_c}\right) \\
& \times \int_0^{\infty} I(\tau > t_s) \exp\left[-V \cdot (\phi_c^*(X) - \phi_c(X))\, I(\tau < t_c)\left(t_c^{\lambda_c} - \tau^{\lambda_c}\right)\right] \\
& \times \exp\left[-V \cdot \left(\phi_s(X) \cdot \tau^{\lambda_s}\right)\right] \lambda_s \tau^{\lambda_s - 1} d\tau \\
& + I(t_c \leq t_s) \exp\left[-V \cdot \left(\phi_c(X) t_c^{\lambda_c} + \phi_s(X) \cdot t_s^{\lambda_s}\right)\right] \\
=\ & I(t_c > t_s) V.\phi_s(X) \exp\left(-V \cdot \phi_c(X) t_c^{\lambda_c}\right) \\
& \times \int_0^{\infty} I(t_s < \tau < t_c) \exp\left[-V \cdot (\phi_c^*(X) - \phi_c(X))\left(t_c^{\lambda_c} - \tau^{\lambda_c}\right)\right] \\
& \times \exp\left[-V \cdot \left(\phi_s(X) \cdot \tau^{\lambda_s}\right)\right] \lambda_s \tau^{\lambda_s - 1} d\tau \\
& + I(t_c > t_s) \exp\left[-V \cdot \left(\phi_c(X) t_c^{\lambda_c} + \phi_s(X) \cdot t_c^{\lambda_s}\right)\right] \\
& + I(t_c \leq t_s) \exp\left[-V \cdot \left(\phi_c(X) t_c^{\lambda_c} + \phi_s(X) \cdot t_s^{\lambda_s}\right)\right]
\end{aligned}
$$

whereas we have the following for controls:

$$S(t_s, t_c | X, V, W = 0) \tag{24}$$
$$= \exp\left[-V \cdot \left(\phi_c(X) \cdot t_c^{\lambda_c} + \phi_s(X) \cdot t_s^{\lambda_s}\right)\right]$$

The integral in (23) can be further "simplified" as:

$$\int_0^\infty I(t_s^{\lambda_s} < \tau^{\lambda_s} < t_c^{\lambda_s})$$
$$\times \left[1 + \omega\left(\phi_c(X)t_c^{\lambda_c} + (\phi_c^*(X) - \phi_c(X))\left(t_c^{\lambda_c} - \left(\tau^{\lambda_s}\right)^{\lambda_c/\lambda_s}\right) + \phi_s(X) \cdot \tau^{\lambda_s}\right)\right]^{-(\alpha+1)} d\tau^{\lambda_s}$$
$$= \int_0^\infty I(t_s^{\lambda_s} < u < t_c^{\lambda_s})$$
$$\times \left[1 + \omega\left(\phi_c^*(X) - \phi_c(X)\right)\left(\left(t_c^{\lambda_s}\right)^{\lambda_c/\lambda_s} - u^{\lambda_c/\lambda_s}\right) + \omega\phi_s(X) \cdot u\right]^{-(\alpha+1)} du$$
$$= \int_p^q \left[1 + \omega\phi_c(X)t_c^{\lambda_c} + \omega\left(\phi_c^*(X) - \phi_c(X)\right)\left(q^r - u^r\right) + \omega\phi_s(X) \cdot u\right]^{-(\alpha+1)} du$$
$$= \frac{1}{a}\left[\omega\phi_s(X)\right]^{-(\alpha+1)} \int_p^q a\left[b + x + c\left(q^r - x^r\right)\right]^{-(a+1)} dx$$

say, where

$$
\begin{aligned}
a &= \alpha \\
b &= \frac{1 + \omega\phi_c(X)t_c^{\lambda_c}}{\omega\phi_s(X)} \\
c &= \frac{\phi_c^*(X) - \phi_c(X)}{\phi_s(X)} \\
p &= t_s^{\lambda_s} \\
q &= t_c^{\lambda_s} \\
r &= \lambda_c/\lambda_s
\end{aligned}
$$

Finally, note that in order for the integral[14]

$$\int_p^q a\left[b + x + c\left(q^r - x^r\right)\right]^{-(a+1)} dx \tag{25}$$

---

[14]The integral (25) is one of the available transformations in the user-defined ML module of EasyReg International, which was used for the empirical work.

to be well-defined, we must require that:

$$a > 0, \ b \geq 0, \ p \geq 0, \ q \geq p, \ r \geq 0, \ \text{and} \ c > -\frac{b+p}{q^r - p^r}. \qquad (26)$$

References

ABBRING, J., AND G. V. DEN BERG (2000a): "The Non-Parametric Identification of Treatment Effects in Duration Analysis," Unpublished Paper, Free University of Amsterdan.

ABBRING, J., G. VAN DEN BERG, AND J. VAN OURS (1997): "The Effect of Unemployment Insurance Sanctions on the Transition Rate from Unemployment to Employment," Working Paper, Tibergen Institute.

ABBRING, J. H., AND G. V. DEN BERG (2000b): "The Non-Parametric Identification of the Mixed Proportional Hazards Competing Risks Model," Manuscript, Free University of Amsterdam.

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 Ofender Recidivism: What Works!," *Criminology*, 34(4), 575 – 607.

HECKMAN, J., R. J. LALONDE, AND J. A. SMITH (1999): *The Economics and Econometrics of Active Labor Market Programs* vol. 3A of *Handbook of Labor Economics*, chap. 31, pp. 1865–2097. Elsevier Science, Amsterdam.

LANCASTER, T. (1990): *The Econometrics Analysis of Transition Data,* Econometric Society Monographs. Cambridge University Press.

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.

MYLKMAN, 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*. The MIT Press, Cambridge, MA, second edn.

SCHMIDT, P., AND A. D. WITTE (1988): *Predicting Recidivism Using Survival Models*, Research in Criminology. Springer-Verlag.

TIMROTS, A. (1985): "An Evaluation of Employment Services Programs for Ex-Offenders," Master's thesis, University of Maryland, College Park.

VAN DEN BERG, G. (1999): "Empirical Inference with Equilibrium Search Models of the Labor Market," *The Economic Journal*, pp. F283 – F306.

——— (2000): *Duration Models: Specification, Identification, and Multiple Durations* vol. V of *Handbook of Econometrics*. North Holland, Amsterdam.

VAN DEN BERG, G., M. LINDEBOOM, AND G. RIDDER (1994): "Attrition in Longitudinal Panel Data and the Empirical Analysis of Dynamic Labour Market Behavior," *Journal of Applied Econometrics*, 9, 421–435.

VAN DEN BERG, G., B. VAN DER KLAAUW, AND J. VAN OURS (1998): "Punitive Sanctions and the Transition Rate from Welfare to Work," Working Paper, Tinbergen Institute.