Dynamic Evaluation of Job Search Assistance

Stephen Kastoryano* Bas van der Klaauw†

July 14, 2015

Abstract

This paper evaluates a job search assistance program for unemployment insurance recipients. The assignment to the program is dynamic. We discuss dynamic treatment effects and identification conditions. In the empirical analysis we use administrative data from a unique institutional environment in which we know the variables determining assignment to treatment. We compare results from different dynamic discrete-time evaluation models and continuous-time duration models. All approaches show that the job search assistance program reduces exit to work, in particular when provided early during the spell of unemployment. We conclude that the discrete-time approach makes less strict parametric assumptions, but the results are sensitive to the choice of control group and the unit of time. [Note: Supplementary STATA and Ox estimation code]

Keywords: treatment evaluation, dynamic enrollment, g-computation, labor market policy.

JEL-code: C22, J64, J68.

*University of Mannheim. E-mail: s.kastoryano@uni-mannheim.de

†VU University Amsterdam, Tinbergen Institute.

Address: Department of Economics, VU University Amsterdam, De Boelelaan 1105, NL–1081 HV Amsterdam, The Netherlands. E-mail: b.vander.klaauw@vu.nl

Thanks to Pierre Koning and seminar participants at the University of Tilburg, University of Zurich, University College London, University of Lausanne, Bristol University, Netherlands Bureau of Economic Policy Analyses, University of Essex, EALE/SOLE-meeting in London, the ESWM in Shanghai, and the IZA/IFAU conference on Labor Market Policy Evaluation 2010 for useful comments and Lukas Hoesch for excellent research assistance.

1 Introduction

Since the 1990s, many countries offer job search assistance to stimulate the exit to work of unemployed workers. Policy makers often consider this to be a necessary requirement in a system with relatively generous benefits. In their survey, Card, Kluve and Weber (2010) stress that job search assistance programs often have relatively good short-run effects. Job search assistance programs are offered frequently in the Netherlands. In fact, the Netherlands is one of four OECD countries spending more than one percent of GDP on active labor market programs (see OECD, 2010). However, empirical evidence on the effectiveness of Dutch active labor market programs is limited, in particular for job search assistance programs. In this paper, we focus on job search assistance for unemployed workers in the Dutch primary education sector. The outcome variable we consider is the exit from unemployment, which is also the key variable of interest to policymakers. Focusing, for example, on wages is less interesting since a majority of the unemployed teachers return to working at primary schools where wages are determined by collective bargaining and are almost a deterministic function of the individual's age (with some extras for managerial responsibilities).

In the Netherlands, the use of randomized social experiments in social insurance schemes is less common than in North-American countries (e.g. Card and Hyslop, 2005, Ham and LaLonde, 1996, LaLonde, 1986). Van den Berg and Van der Klaauw (2006) describe the most recent experiment conducted in 1998/99. In our paper we deal with non-experimental data. Non-experimental data make the empirical evaluation non-trivial. First, the evaluation of job search assistance suffers from the usual endogeneity problem that participation might be related to (unobserved) individual characteristics. Second, job search assistance often does not start immediately after an individual enters unemployment which means that some individuals leave unemployment before the moment they should enter job search assistance. Furthermore, constructing a counterfactual is non-trivial if the start of the program differs between individuals, and all individuals could eventually enter job search assistance. This problem is known as dynamic selection (see Abbring and Heckman, 2007; for a survey on dynamic treatment evaluation).

Ham and LaLonde (1996) and Eberheim, Ham and LaLonde (1997) provide early contributions to the empirical evaluation of an active labor market program in a dynamic setting. Eberwein, Ham and LaLonde (1997) obtain causal effects of treatment on the exit rate to work in a bivariate duration model with unobserved heterogeneity using an instrumental variables approach. Identification then is obtained assuming there exists a sequentially randomized intention to treat which can be used to instrument actual treatment. We do not have such instruments at our disposal. However, the advantage of our setting is that we know which variables determine assignment to treatment and that job seekers do not know precisely when they are assigned to the job search assistance program.

Even though we only evaluate a single treatment, the effects can differ between individuals. Not only because individuals are heterogeneous, but also because the impact of the program can depend on the moment of starting the job search assistance. We investigate to what extent the timing of entering job search assistance influences its effectiveness. This information is important for the targeting efficiency of the program. Since the Dutch government is facing substantial budget cuts, the issue of which unemployed workers to assign to labor market programs becomes more important.

In our empirical analyses, we use administrative data from a unique institutional environment in which the assignment to job search assistance is very clearly described. The participation in job search assistance depends only on a limited set of observable characteristics which we observe. Since we know the institutional environment and variables determining selection into the program we can credibly justify identifying assumptions in the different evaluation models. The results provide valuable insights in the performance (in a real life setting) of some estimators for dynamic policy evaluation.

In the US, interventions during unemployment often start at a fixed moment. Black, Smith, Berger and Noel (2003), for instance, study training services starting after two weeks of unemployment. A substantial share of the econometric methodology therefore focuses on static treatment evaluation (e.g. Imbens and Wooldridge, 2009; for a recent survey). However, in many European countries the timing of entry into labor market programs varies between individuals. Other than in the Netherlands, this is the case in Sweden (e.g. Fredriksson and Johansson, 2008; and Sianesi, 2004), Switzerland (e.g. Gerfin and Lechner, 2002; and Lalive, Van Ours and Zweimüller, 2008), France (e.g. Crépon, Ferracci, Jolivet and Van den Berg, 2010), and Germany (e.g. Lechner and Wunsch, 2009).

A relatively large literature attempts to fit such dynamic settings into the standard potential outcome framework (e.g. Gerfin and Lechner, 2002; and Sianesi, 2004). Sianesi (2004) discusses the complications of finding a suitable control group in the case where there is ongoing entrance into the program. In such cases, constructing the control group for individuals who enter the program at one moment is complicated by the fact that other individuals enter the program later. Considering for the control group only those individuals who are observed not to have received treatment implies conditioning on future outcomes. Fredriksson and Johansson (2008) argue that treatment evaluation estimators will be biased if the future outcomes depend on dynamic selection and dynamic selection is not accounted for in the control group.

This paper fits within the recently growing literature on dynamic treatment evaluation, surveyed by Abbring and Heckman (2007). Our contribution is not only empirical, we also intend to make some methodological contributions. We discuss the implementation of both discrete-time and continuous-time methods for dynamic treatment evaluation. Our discrete-time approaches include dynamic matching approaches (Gerfin and Lechner 2002; Sianesi, 2004; Fredriksson and Johansson, 2008) and an easy form of the g-computation formula introduced by Robins (1986) in the bio-statistical literature (see also Robins, 1997;

and Gill and Robins, 2001). Methods from the bio-statistical literature are still relatively unknown in the economic literature but offer more flexibility in constructing outcome distributions than usual reduced-form dynamic discrete-time models. We discuss the relevance of the assumptions underlying the discrete-time g-computation formula in an economic setting. We also compare these results to those from a continuous-time mixed proportional hazard model (Abbring and Van den Berg, 2003).

One key identifying assumption for discrete-time models is the sequential unconfoundedness assumption. This assumption is often justified from the richness of the data. Gerfin and Lechner (2002) and Sianesi (2004), for example, argue that information on past labor market outcomes and subjective assessments of labor market prospects justify the sequential unconfoundedness assumption. Lalive, Van Ours and Zweimüller (2008) show that even if such information is available, applying timing-of-events estimation and propensity score matching estimation give substantially different results. Unlike our institutional setting, in their setting it is unclear whether the conditional independence assumption is valid.

The paper is organized as follows. In Section 2 we provide details about the relevant unemployment insurance scheme, and the job search assistance program. Section 3 presents the data. In Section 4 we provide a general framework for dynamic treatment evaluation. Section 5 discusses our estimation procedure and Section 6 presents the estimation results. Finally, Section 7 concludes.

2 Institutional setting

2.1 Unemployment insurance for the primary education sector

Our data concern former employees of Dutch primary education institutions who are entitled to collecting unemployment insurance benefits. Primary education institutions, like all public sector institutions, must bear their own unemployment insurance risk. However, because primary education institutions are relatively small, they were forced in 1996 to participate in a sector fund, called the Participation Fund. This fund is responsible for collecting premiums, and paying unemployment insurance benefits.

Unemployed workers from the primary education sector have the same entitlement rules and obligations as unemployed workers from the private sector. Their benefits are, however, more generous both in terms of level and entitlement period. All individuals below age 65 who worked at least 26 weeks of the 36 weeks prior to becoming unemployed are entitled to collecting unemployment insurance benefits. Furthermore, a worker should have lost at least five working hours per week or more than 50% of their weekly working hours (if less than 10). Finally, the job loss should not be voluntary, and the individual should not be held responsible for the job loss.

Each unemployed worker receives unemployment insurance benefits for at least three months. If an unemployed worker worked at least 52 days during four out of the past five calendar years ('year'-condition), the entitlement period is extended to six months. For each additional year of employment (so beyond four years) the entitlement period for unemployment insurance benefits is extended by one month. For an entitlement period of one year, the unemployed worker must have worked for at least ten years. For the maximum entitlement period of 38 months, 36 years of work is required. During the first year, the benefits level is 78% of the last wage (capped at 167.70 euro per day). After that, the benefits level decreases to 70% of this last wage.

After the usual benefits entitlement period ends, an individual may be entitled to extended benefits at 70% of the last wage. The duration of the extended benefits depends on age and work experience. Individuals below age 40 and those with less than five years of work experience do not receive extended benefits. A 40-year old individual with five years of work experience receives one additional year of benefits, while a 51-year old with more than ten years of work experience receives extended benefits until reaching the retirement age of 65.

Benefit recipients have the obligation to actively search for work, and to accept suitable job offers. Furthermore, they should provide all necessary information to the Participation Fund, and keep them informed about possible changes in their situation (e.g. vacation, sickness, pregnancy, etc.). If the individuals fail to comply to these rules, a sanction can be applied which leads to a temporary reduction of the benefits level.

During our observation period, the unemployment rate in the primary education sector was about 2% compared to 4% in the private sector. The main reason for the lower unemployment rate is a much lower inflow. The outflow from unemployment in the primary education sector is comparable to that of the private sector. There are compositional differences between unemployed workers in the primary education sector and the private sector. About 80% of the workers in primary education are women, and the average age is somewhat higher than in the private sector.

2.2 The job search assistance program

Since July 2005, institutions in the public sector are also responsible for reintegrating their former employees. This implies that the Participation Fund became responsible for financing and organizing active labor market programs. These activities fall into two categories. First, a regular program in which the majority of the benefit recipients participate. This program focusses on job applications, but can also include some vocational training. Second, a short job search assistance program focussing on networking skills in addition to job application training. These programs are not specific to the primary education sector, unemployed workers from the private sector also participate via the nationwide unemployment insurance administration.

Unemployed workers under age 60 are obliged to participate in these programs if these are offered to them. Individuals who refuse to participate will be sanctioned with a substantial reduction of their benefits. Participation in a program does not affect the entitlement to benefits, i.e. the benefit entitlement period is not extended and individuals do not get additional benefits while being in the job search assistance program. Most individuals aim at finding new work again in the primary education sector, but about one-third of the observed exits are towards employment outside this sector.

The program is offered to individuals who receive benefits for at least eight hours per week, and with an entitlement period exceeding three months. Individuals with less than 13 months entitlement at the moment of entering the program are assigned to the short program. Individuals with a longer entitlement period enter the regular program. The timing of assignment to the program differs depending on an age criteria. All individuals above age 50 (at the first day of unemployment) and (low-skilled) individuals who were previously employed in a subsidized job, are supposed to enter the job search assistance program immediately after starting collecting benefits. Individuals under age 50 and who are not low-skilled, enter the program only after six months of unemployment.

Only 8% of all job search assistance programs offered are short programs. These services last three months and focus on presentation, writing a vitae and application letter, networking and efficient job search. The remaining 92% of the job search assistance programs offered are regular programs. In the empirical analyses we do not distinguish between both programs. The programs are offered at 11 locations providing all the same program. Once invited the benefit recipients can choose the location but 75% of the individuals accept the default. The remaining 25% almost always opt for the location nearest to their home.

The regular program starts with an intake interview to determine the required activities. These range from improving language skills, providing psychological support, providing short vocational class, and offering the type of job search assistance services also included in the short program. The training takes place both in individual and group meetings. The intensity of the meetings depends upon the needs of the individual. The first weeks are often more intense, with two to three meetings per week with training officers. The total time spent in these meetings is about one full working day per week. After this period, the participants usually visit the training center once a week or every other week for a few hours. During this later stage, participants receive weekly assignments to be discussed in the weekly meetings. The general goal is that after two months of participating in the program individuals should start making successful job applications. However, participation in the program does not lower the job search requirements. While in the job search assistance program, unemployed workers must comply to the same minimum job applications requirements as when not being in the program. The job search assistance program should not last longer than one year, and individuals who start a new job during the program are offered to finish the program while working. The cost of the short job search assistance program is 500 euro per individual entering the program. The cost of the regular job search assistance program is 4000 euro for individuals above age

50 and for low-skilled individuals, and 3750 euro for individuals below age 50.[1]

The Participation Fund does not assign benefit recipients directly to programs, but outsources this task to a separate firm. This firm never has any personal contact with unemployed workers and receives only a limited amount of information when assigning them to treatment. The information consists of the social security number, gender, age, an indication for being low-skilled (i.e. previously in a subsidized job), entitlement duration to benefits, number of weekly hours of collecting benefits, and an indicator code for the previous employer.[2] Two weeks prior to the start of the program the individual receives a letter explaining that she should enroll in a program. This letter also offers individuals to select one of the 11 locations.

In practice the policy guidelines concerning the timing of entering job search assistance were not followed strictly. This was due to administrative and communication issues between the Participation Fund and the external firm.[3] There are cases where records got lost, where information was provided too late, and where notification letters were never sent. As we will show in the next section, this creates substantial variation in the assignment of the program. And, since the external firm never had any contact with benefit recipients, the variation in program assignment should be exogenous conditional on observed individual characteristics. We exploit the latter in the empirical analyses.

3 Data

In the empirical analysis we use administrative data from the Participation Fund. Our data concern all former employees from primary education institutions who started collecting unemployment insurance benefits between August 1, 2006 and April 1, 2008. Individuals are followed until their benefits entitlement ends (due to finding work or having exhausted their entitlement period) or until March 12, 2009. From the data we only consider those individuals who started collecting benefits within 30 days after being laid-off. According to the job search assistance program criteria, we leave out individuals who claimed benefits for less than eight hours per week. We also exclude individuals above age 60 since for them participation in the job search assistance program is voluntary.

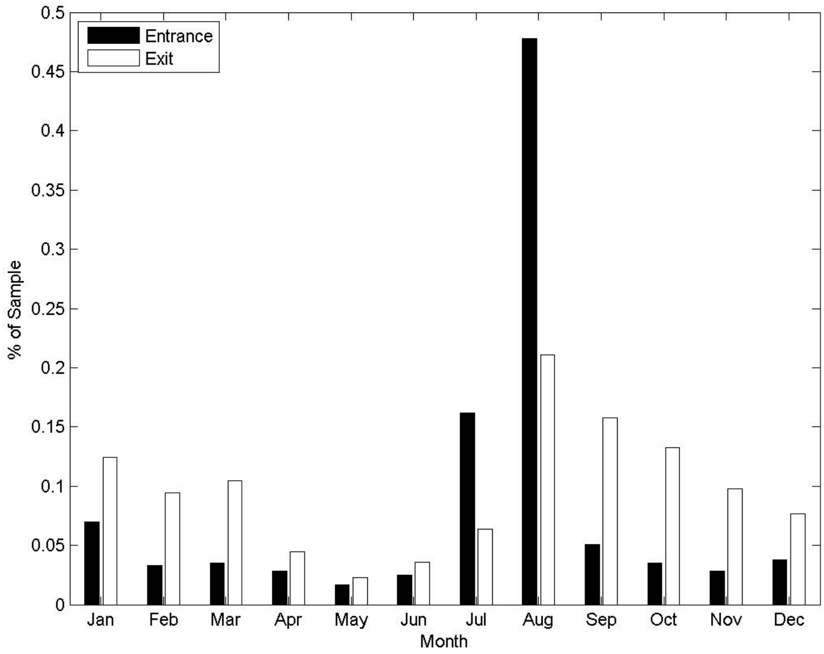
From the data we drop three individuals who very often entered and exited unemployment during the observation period. We exclude 43 observations for which the date of entering the job search assistance program was unknown or prior to becoming unemployed. The latter might occur if the individual was still in the program from an earlier

[1]Some commercial active labor market programs have some pay-for-performance scheme, which yield more incentives.

[2]The policy is to avoid having individuals previously employed at the same institution in the same meeting groups.

[3]In the Netherlands, all individuals applying for unemployment insurance benefits should apply at the nationwide UI administration. This administration forwards files of workers from the primary education sector to the Participation Fund, which already causes a delay ranging from a few days to a few weeks.

Figure 1: Seasonal variation in entry into and exit from unemployment.



unemployment spell. Finally, we exclude 37 observations with an hourly wage in the previous job below three euro, which is far below the legally binding minimum wage.

The data contain 3064 individuals for which we only consider the first observed unemployment spell. Over 60% of the individuals are entitled to benefits for more than one year, and 40% have an entitlement period exceeding two years. As can be seen from Figure 1 almost 50% of the inflow occurs in August, which is the start of a new school year. The outflow is much more spread over the year, although there is a decreasing trend over the school year. Figure 2a provides a Kaplan-Meier estimate for the exit to work. The median unemployment duration is about 21 weeks. Of the 3064 individuals, 862 entered the regular job search assistance program and 78 the short program. Figure 2b shows the Kaplan-Meier estimate for entering a program. In the figure we distinguish two groups, those who should enter a program immediately (either older than 50 or low-skilled), and those who should enter after six months of unemployment (below 50 and not low-skilled). The figure clearly shows that the latter group enters the program, on average, later during the unemployment spell. Nevertheless, within each group there is still substantial variation in the moment of entering. This confirms that the external firm did not manage to correctly implement the rules for program assignment.

The data contain a limited set of individual characteristics. In Table 1 we provide some descriptive statistics. We distinguish between individuals who participated in a program during unemployment (participants) and those who did not (nonparticipants). The data contain the same individual characteristics as provided to the external firm

Figure 2: Kaplan-Meier estimates.

(a) Exit to work. (b) Entering job search assistance.

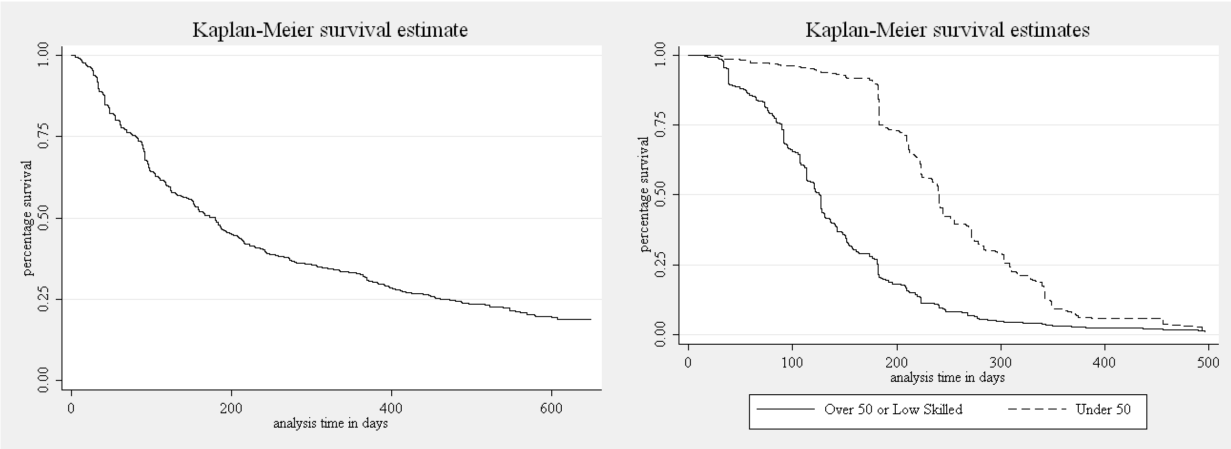


Table 1: Descriptive statistics.

	Program participants	**Non-participants**
Number of observations	940	2124
Median unemployment duration (in days)	369	96
Median duration to program start (in days)	156	
Unemployment hours per week	29.9	26.7
Benefits level (hourly)	€12.8	€10.4
Female	64%	85%
Age 20-35	9%	59%
Age 35-50	46%	29%
Age 50-65	45%	12%
Low-skilled	34%	4%

who assigned the program. The participants are, on average, unemployed for more hours per week, and have a higher benefits level. This might be the direct consequence of the difference in age structure. Older workers are more likely to participate in a program, which follows from the program assignment policy. Of course, the different composition between the participants and the nonparticipants is not only the result of the assignment mechanism and the implementation of the external firm. Dynamic selection also plays an important role. Those individuals with adverse characteristics have, on average, longer unemployment durations and are thus more likely to have entered the job search assistance at some stage.

4 Model for dynamic treatment evaluation

In this section we briefly discuss a framework for dynamic treatment evaluation. Our discussion draws from the discussion provided by Abbring and Heckman (2007) and from the bio-statistical literature on dynamic treatment effects (e.g. Robins, 1997; Gill and Robins, 2001; Lok, Gill, Van der Vaart and Robins, 2004; and Lok, Hernán and Robins, 2007). We mainly focus on issues relevant for estimating dynamic treatment effects in our policy setting.

4.1 Theoretical framework

Consider the case where we observe for each individual the duration $T > 0$ of unemployment. We define the binary variable Y_t as indicator for being unemployed ($Y_t = 0$) or employed ($Y_t = 1$) after t periods, so $Y_t = I(T \leq t)$. Since the outcome variable describes the survival in unemployment, $\mathrm{E}[Y_t] = 1 - \Pr(T > t)$, where $\Pr(T > t)$ is the survivor function.

Individuals can receive a single treatment only once during the period of unemployment. This setting is discussed by Abbring (2003) and Abbring and Van den Berg (2003). Robins (1997) presents a framework where a binary treatment choice is made in each period while Gill and Robins (2001) consider changing the treatment intensity over time. All individuals in our data are eligible for entering treatment. However, the timing of entering treatment differs between individuals. This is partly systematic because unemployed workers over age 50 and low-skilled individuals should enter the job search assistance immediately, while unemployed workers under age 50 should wait six months. But there is also (nonsystematic) variation in the moment of entering the program.

Robins (1997) and Gill and Robins (2001) allow the treatment decision in each time period to depend on earlier outcomes. They define a treatment regime as the set of rules determining future treatment in the population.In our setting individuals can only enter the program when being unemployed and they might leave unemployment before actually starting treatment. Let $S > 0$ be a random variable describing the elapsed unemployment duration at the moment of entering the job search assistance program. Young, Cain, Robins, O'Reilly and Hernán (2011) refer to this type of treatment assignment as a randomized static treatment plan. In a randomized static treatment plan, treatment at s does not depend causally on previous outcomes or time-varying intermediate individual characteristics. In our case only the yet untreated survivors in unemployment can enter the job search assistance program. This suggests that treatment assignment at time k depends on $Y_k = 0$. However, one may argue that the moment of treatment $S = s$ is assigned at the start of unemployment $t = 0$ and treatment is only realized when the individual's unemployment duration T exceeds s.

Gill and Robins (2001) suggest building a counterfactual space on top of the factual outcomes. They argue that whatever sense of counterfactuals exists or does not exist, it is harmless to pretend that the counterfactuals do exist. Let $Y^*_{1,t}(s)$ denote the potential unemployment status after t periods had the individual been treated at s under a given policy regime. Usually, a consistency assumption is made to link potential outcomes to observed outcomes. This assumption implies that if we observe $S = s$, the random variable describing the observed outcome Y_t equals the potential outcome $Y^*_{1t}(s)$. Even though we only consider a single treatment, it may have different effects when initialized at different moments s. Ideally, we would define the potential untreated outcome $Y^*_{0,t}$ as $\lim_{s\to\infty} Y^*_{1,t}(s)$. These are the outcomes if the individual will not enter the job search assistance program. This allows us to define the average treatment effect

$$\Delta_{\mathrm{ATE}}(t, s) = \mathrm{E}[Y^*_{1,t}(s) - Y^*_{0,t}]$$

Knowledge about $\Delta_{\mathrm{ATE}}(t, s)$ for all values of t and s can, for example, be used for a costs-benefits analysis or to estimate the expected duration of unemployment for different moments of entering the job search program. Determining the optimal timing of the intervention is more complicated because optimizing the timing of the intervention changes the treatment regime and can change the effect of entering treatment at a fixed s. Since we only observe data from a single treatment regime, we can not evaluate the effect of the treatment regime or compare the effects of treatment under alternative treatment regimes.

In a typical data set, there are a few complications. Data usually describe a limited period so they are not informative on the right tail of the distribution of unemployment spells. Furthermore, most data sets do not include the moment of entering the program if the worker left unemployment without having participated in the program. In such cases S is a latent variable and we only know that it exceeds the observed unemployment duration. The untreated outcome $Y^*_{0,t}$ is, therefore, a variable that cannot be identified without additional assumptions.

To identify the counterfactual outcomes, we adopt the so-called no-anticipation assumption described by Abbring and Van den Berg (2003). This assumption states that treatment participation only affects later outcomes

$$Y^*_{1,t}(s) = Y^*_{1,t}(s') \qquad \text{if } s \neq s' \qquad \forall t < s, s'$$

Abbring and Heckman (2007) refer to this assumption as no causal dependence of outcomes on future treatments and this assumption is also explicitly made by Fredriksson and Johansson (2008). Gill and Robins (2001) refer to the no-anticipation assumption as harmless in epidemiological applications. In economics applications, this assumption needs to be justified since agents are often considered to be forward looking and current outcomes are the result of dynamic optimization. Because of the no-anticipation assumption,

$$Y^*_{0,t} = Y^*_{1,t}(s) \qquad \forall s > t$$

In our application, no-anticipation rules out that prior to the actual start of treatment unemployed workers already change their job search behavior in response to being assigned to treatment. This rules out threat effects as, for example, measured by Black, Smith, Berger and Noel (2003). They find that unemployed workers are more likely to leave the benefits program once they have been informed about the actual start of a job search assistance program. A similar result is found by Crépon, Ferracci, Jolivet and Van den Berg (2010).

Imposing no-anticipation does not rule out that individuals know they are exposed to the risk of participating in treatment and may therefore behave differently than in a system in which treatment is absent (Heckman and Navarro, 2007). Justifying the no-anticipation assumption requires knowledge about the unemployed worker's information set prior to actually starting the treatment. In our case, the unemployed workers are informed (by letter) two weeks prior to the start of the job search assistance program. Our data contain some information about these letters which we exploit to justify the no-anticipation assumption (see the next subsection).

Most microeconometric applications aim at estimating the average treatment effect on the surviving treated, which is defined as,

$$\Delta_{\text{ATEST}}(t,s) = \mathrm{E}\left[Y^*_{1,t}(s) - Y^*_{0,t}|S=s, Y_s=0\right] \qquad \text{with } t > s$$

This treatment effect denotes the effect of providing treatment at s on exit to work between s and t for those who survived in unemployment for s periods. This is the ex-post effect of the treatment, so the effect of actually participating in the treatment on future outcomes.

To estimate $\Delta_{\text{ATEST}}(t,s)$, unemployed workers treated at s should be compared to similar unemployed workers who (possibly) receive treatment after t. The main complication is that it is unclear which individuals qualify for the control group. There is, of course, the selection problem if treatment is not assigned randomly. However, an additional problem is that in a setting with ongoing entry in treatment it is not possible to know whether untreated individuals who left unemployment between s and t would have received treatment before or after t. It is unclear how to deal with such observations. Gerfin and Lechner (2002) include these observations in the control group, but exclude individuals who are observed to have received treatment between s and t. This causes a bias towards shorter unemployment spells in the control group, and treatment effects will be underestimated. Ignoring from the control group all individuals who exit between s and t will clearly not either solve the issue.

Sianesi (2004) suggests for the control group all surviving individuals who are not treated at s. This implies that the treatment effect changes to

$$\Delta^*_{\text{ATEST}}(t,s) = \mathrm{E}\left[Y^*_{1,t}(s) - Y^*_{1,t}(S>s)|S=s, Y_s=0\right] \qquad \text{with } t > s$$

where $Y^*_{1,t}(S>s)$ is the potential employment outcome at t for an unemployed worker not treated before or at s. This treatment effect describes the effect of entering treatment at s compared to possibly entering treatment at some later moment. The usefulness of this treatment effect is limited, mainly because the counterfactual outcomes and also the treatment effect depend on the future entry process into treatment and effect of the treatment in the control group. A costs-benefits analysis, for example, is not straightforward since it is unclear when individuals in the control group receive treatment.

Both approaches mentioned above to construct counterfactuals are mainly driven by the requirement to fit the evaluation problem within the standard (static) potential outcome model. Dynamic techniques can deal with such data problems more flexibly. In the next section, we discuss both a discrete-time and a continuous-time approach dealing with these issues. In both approaches, we exploit that selection is on observables. However, conditional on observed variables, both approaches make the no-anticipation assumption discussed above. Therefore, we first justify this crucial assumption.

4.2 Justifying no-anticipation

In the previous subsection we introduced the no-anticipation assumption. If unemployed workers receive information about the timing of entry in job search assistance far before the actual start, they might take this into account in their current job search decisions. Crépon, Ferracci, Jolivet and Van den Berg (2010) use notifications to test for anticipation of training programs. They find strong effects of the notifications already before the start of the training program. In their setting the average time between notification and entry in the program is almost three months.

Our data also contain some information on invitation letters for the job search assistance program, which should be sent about two weeks prior to the start of job search assistance. However, this information is very incomplete. Letters are only recorded since April 2008, so no information is available on the first two years of the observation period. There is also no guarantee that for the later period the information on the letters is complete. In total, we observe that 279 letters were sent. We observe only four individuals who left unemployment in the two weeks prior to receiving the letter, but no one in the short period after receiving the letter. Furthermore, the data show that in almost all cases only 14 to 20 days elapsed between the sending of a letter and the start of a job search assistance program. This provides for our setting evidence in favor of the no-anticipation assumption.

The assumption of no-anticipation does not mean that individuals do not know about the assignment rules for the job search assistance program. Unemployed workers may be informed about the assignment rules. For example, an individual above age 50 may know that she should enter the program as soon as possible. However, she can not precisely choose the timing of reemployment given exact knowledge of the future timing of entering the training. This also implies that individuals cannot manipulate their assignment to the program. Given the assignment mechanism through the external agency assigning job search assistance, it is unlikely that individuals can either manipulate or obtain prior knowledge about their actual assignment date. Unemployed workers do not know about the existence of the external firm, and the external firm only receives very limited information about each unemployed worker.

5 Empirical analysis

We consider two possible approaches to estimate the treatment effects defined in the previous section. First, a discrete-time model based on the g-computation algorithm (Robins, 1997).[4] Second, a continuous-time duration model as discussed by Ridder (1986). We first describe both methods and in Subsection 6 we discuss our estimation results and compare both approaches.

5.1 Discrete-time approach

In our discrete-time approach one unit of time is denoted by κ. This defines the interval cut-off points $\tau_{k+1} = \tau_k + \kappa$ with $\tau_0 = 0$ the moment at which the worker's unemployment spell starts. The indicator $D_k = I(S \leq \tau_k)$ describes the treatment status in period k which implies that $D_k = 0$ if $D_{k+1} = 0$ and $D_{k+1} = 1$ if $D_k = 1$. We define the treatment sequences $\bar{D}_k(s) = \{D_0 = 0, \ldots, D_{s-1} = 0, D_s = 1, \ldots, D_k = 1\}$ for treatment at $\tau_{s-1} < S \leq \tau_s$.

Furthermore, the potential outcomes $Y^*_{1,k}(s)$ describe the employment status at τ_k had the worker entered the job search assistance program in the interval $(\tau_{s-1}, \tau_s]$, so if she follows the treatment sequences $\bar{D}_k(s)$. We observe the outcomes $Y_k = I(T \leq \tau_k)$. These outcome variables may be latent when the unemployment spell is right censored prior to τ_k. We assume that right censoring is exogenous.

In the remainder of this section we maintain the no-anticipation assumption and the consistency assumption discussed in the previous section. Let X denote the set of observed characteristics which include all individual characteristics provided to the external firm to assign unemployed workers to the job search assistance program. Since our data contain all variables used for the assignment to training we can make a strong case that, conditional on X, there are no unmeasured confounders jointly determining assignment to the job search assistance program and future potential outcomes,

$$Y^*_{1,k'}(s) \perp\!\!\!\perp D_k | D_{k-1} = 0, Y_k = 0, X = x \qquad \forall k > 0 \text{ and } k' > k \text{ and } s \geq k$$

This unconfoundedness assumption says that potential future outcomes are independent of treatment assignment in period k among the survivors after k periods with observed variables $X = x$ who were untreated up until period $k-1$. This is a dynamic version of a conditional independence assumption. In our formulation, by conditioning on $D_{k-1} = 0$ and $Y_k = 0$ we assume training in interval τ_k only influences the job arrival rate and the job acceptance rate starting interval τ_{k+1}. This imposes that job applications take some

[4]Robins (1997) defines a treatment regime g as a vector of assignment rules mapping the history of treatment and observed variables to a current treatment.

time to process. However, it excludes that workers change their job acceptance rate in the interval τ_k due to entering the training program in interval τ_k.[5] Our assumption is slightly less strict than in Robins (1997), but combined with the no-anticipation assumption it has the same implications.

In addition, we follow Robins (1997) and Lok, Gill, Van der Vaart and Robins (2004) by imposing that treatment sequences are evaluable,

$$0 < \Pr(D_k = 1|D_{k-1} = 0, Y_k = 0, X = x) < 1 \quad \text{if} \quad \Pr(D_{k-1} = 0, Y_k = 0|X = x) > 0 \quad \forall k > 0 \quad \text{and} \quad x$$

This common support assumption guarantees that at any moment exposure to the job search assistance program is not deterministically allocated among untreated survivors. Recall from Figure 2 that for both assignment rules to treatment, there are indeed individuals entering the program at each unemployment duration.

In our discrete-time notation, we can rewrite the treatment effects of interest from the previous section as

$$\begin{aligned}
\Delta_{\text{ATEST}}(\tau_k, \tau_s) &= \mathrm{E}\left[Y^*_{1,k}(s) - Y^*_{0,k}|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0\right] \\
&= \int_x \Big[\mathrm{E}[Y^*_{1,k}(s)|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0, X = x] f_X(x|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0)dx \\
&\quad - \int_x \mathrm{E}[Y^*_{0,k}|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0, X = x] f_X(x|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0)dx
\end{aligned}$$

where $F_X(x|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0) = F_X(x|\tau_{s-1} < S \leq \tau_s, Y_s = 0)$ is the distribution of observed covariates for the non-treated survivors who enter the program in the interval $\tau_{s-1} < S \leq \tau_s$.

We can derive an expression for the conditional expectation of the potential treated outcomes as follows,

$$\begin{aligned}
&\mathrm{E}[Y^*_{1k}(s)|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0, X = x] \\
&\quad = 1 - \Pr(Y^*_{1k}(s) = 0|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0, X = x) \\
&\quad = 1 - \Pr(Y_k = 0|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0, X = x) \\
&\quad = 1 - \prod_{j=s+1}^{k} \Pr(Y_j = 0|\bar{D}_{j-1}(s) = \bar{d}_{j-1}(s), Y_{j-1} = 0, X = x)
\end{aligned}$$

This expression assumes there is no right censoring of unemployment durations. In case of right censoring, then the above must condition each transition probability on the non-censored at period j and add an additional unconfoundedness assumption.[6]

[5] More formally, we assume $\Pr(Y_s = 0|D_s = 1, D_{s-1} = 0, Y_{s-1} = 0, X = x) = \Pr(Y_s = 0|D_s = 0, D_{s-1} = 0, Y_{s-1} = 0, X = x)$.

[6] More formally, define C_k as an indicator taking value 1 if an observation is censored at or before period τ_k. In our counterfactuals we impose the missing completely at random assumption which is

$$(Y^*_{1,k'}(s), D_k) \perp\!\!\!\perp C_k|D_{k-1} = 0, Y_k = 0, X = x \qquad \forall k > 0 \text{ and } k' > k \text{ and } s \geq k$$

Deriving the expression for the counterfactual outcomes is slightly more complicated since we have to deal with enrollment in the job search assistance program between τ_s and τ_{k-1}. The expectation of the counterfactual outcomes equals

$$
\begin{aligned}
\mathrm{E}[Y^*_{0k}|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0, X = x]& \\
&= 1 - \Pr(Y^*_{0k} = 0|\bar{D}_k(s) = \bar{d}_k(s), Y_s = 0, X = x) \\
&= 1 - \Pr(Y^*_{0k} = 0|D_s = 0, Y_s = 0, X = x) \\
&= 1 - \prod_{j=s+1}^{k} \Pr(Y_j = 0|D_{j-1} = 0, Y_{j-1} = 0, X = x)
\end{aligned}
$$

So now the treatment effect of interest $\Delta_{\text{ATEST}}(\tau_k, \tau_s)$ can be written in terms of a series of transition probabilities. The second equality exploits the no-unmeasured confounders assumption and in the third step the link is made between the potential outcomes and the observed outcomes. The last step also relies again on the no-unmeasured confounders and no-anticipation assumptions. This formula evaluates all possible paths leading to the observed outcome Y_k and starting from not observing program enrollment before τ_s. In Appendix A we also derive the propensity score weighting version of the g-computation formula.

In the estimation we construct a matched control sample, such that in this sample $f^m_X(x|S > \tau_s, Y_s) = f_X(x|\tau_{s-1} < S \le \tau_s, Y_s)$.[7] As controls we take all untreated individuals whose propensity score differs less than 0.025 from the propensity score of the treated individual which is estimated with a Logit specification. If there are multiple controls for a single treated individual we weight the observations using a Gaussian kernel with bandwidth 0.005.[8] Since we observe all variables determining program participation, matching adjusts for any ex-ante differences in exit rates prior to period s.

Matching on covariates also allows us to estimate ex-post exit probabilities separately for the treated and weighted matched controls using Kaplan-Meier estimators. However, we still face the problem of dynamic selection due to the simultaneity of treatment and exit within an interval for the control group. To adjust for this dynamic selection bias we weigh each control unit at period $j+1$ by the probability of program enrollment in period j $\Pr(D_j = 0|D_{j-1} = 0, Y_j = 0, X = x)$ which we estimate with a Logit specification.This follows Ham and Rea (1987) who also used Logit specifications for transition probabilities. We include as covariates the worker's gender, age, age^2, logarithm of pre-unemployment wage, the duration until exhaustion of benefits, an indicator for becoming unemployed in July or August, an indicator for being above 50 years of age, and an indicator for

which means that for the survivors up to period τ_k who are untreated up until period τ_{k-1}, we can ignore (leave out) observations censored in the interval τ_k. The missing completely at random assumption which we assume in estimation excludes the $X = x$ from the conditioning set.

[7] Computer code in STATA, and Ox 6.10 for all dynamic evaluation methods discussed in this paper can be found at skastoryano.com. The code includes options for matching based on the propensity score, nearest neighbor matching, and propensity score weighting.

[8] Most of our results are robust against the choice of matching algorithm.

coming from a low-skilled job. To compute standard errors we apply subsampling on the untreated survivor population (Politis and Romano, 1994).[9] We additionally smooth the estimates over a period of 10 days around the treatment time.

5.2 Continuous-time approach

The previous approach imposes discrete time intervals even though we observe unemployment durations at a daily level. Another approach is to allow for a continuous-time model specification. Consider an individual collecting benefits for t units of time. We assume that the exit rate can be characterized by observed characteristics X, unobserved characteristics V, the elapsed unemployment duration T itself, and a variable $I(s < t)$ indicating whether the individual already started participating in the job search assistance program, where s is the moment at which an individual enters job search assistance.

The exit rate from unemployment at t conditional on $X = x$, $V = v$ and $S = s$ is denoted by $\theta(t|x, v, s)$, and follows the mixed proportional hazard specification

$$\theta(t|x, v, s) = \lambda(t) \exp(x'\beta + \delta \cdot I(s < t) + v)$$

The entry rate into job search assistance at t conditional on observed and unobserved characteristics $X = x$ and $V_s = v_s$ is similarly given by

$$\theta_s(t|x, v_s) = \lambda_s(t) \exp(x'\beta_s + v_s)$$

We parameterize the duration dependence functions $\lambda(t)$ and $\lambda_s(t)$ as piecewise constant. In these specifications, the unobserved heterogeneity terms are random effects uncorrelated to X. It is also possible to specify several discrete mass points for the distributions of the unobserved heterogeneity terms of $G(v)$ and $G_s(v_s)$.[10] This specification is flexible but including many time intervals or mass point locations requires many observations. Maximum likelihood is used to estimate the model parameters.

The identification of this continuous-time model requires the no-anticipation assumption and the assumption of no unmeasured confounders. Furthermore, this model assumes multiplicative effects of covariates and the treatment effect on the hazard rates. Thus, while the continuous time model does not require specifying the length of time intervals, it imposes a strong parametric restriction. In the next section we will see that this parametric restriction influences estimation results.

[9]Since there is no clear rule guiding the selection of subsample size we fix the number of subsamples $b = int(n^{95/100})$ which will vary depending on the number of untreated survivors n. This choice of subsamples is more conservative than Lalive, Van Ours and Zweimüller (2008) who set $b = int(n^{99/100})$. We find that the number of subsamples affects the estimated standard errors.

[10]Kastoryano and Van der Klaauw (2011) provide empirical evidence that additional mass points do not affect estimates in the evaluation of this data. This also supports the validity of the no unmeasured confounders assumption.

Since the continuous-time model fully specifies the hazard rates to the job search program and out of unemployment we can obtain an estimate for the treatment effect $\Delta_{\text{ATEST}}(t,s)$. We first define for unemployed worker i with observed characteristics x_i,

$$\mathrm{E}[Y^*_{1,t}(s) - Y^*_{0,t}|Y_s = 0; x_i, v] = \frac{\exp(-\int_0^t \theta(z|x_i,t,v)dz) - \exp(-\int_0^t \theta(z|x_i,s,v)dz)}{\exp(-\int_0^s \theta(z|x_i,s,v)dz)}$$

To translate this into the average treatment effect on the treated $\Delta_{\text{ATEST}}(t,s)$, we then condition on the rate of receiving treatment after s periods. This allows us to estimate the average treatment effect on the surviving treated as

$$\Delta_{\text{ATEST}}(t,s) = \frac{\int_v \int_{v_s} \sum_i f(s|x_i,v,v_s)\mathrm{E}[Y^*_{1,t}(s) - Y^*_{0,t}|Y_s = 0; x_i, v]G_s(v_s)G(v)}{\int_v \int_{v_s} \sum_i f(s|x_i,v,v_s)G_s(v_s)G(v)}$$

where $f(s|x_i,v,v_s) = \theta_s(s|x_i,v_s)\exp(-\int_0^s \theta(z|x_i,v,s)dz)\exp(-\int_0^s \theta_s(z|x_i,v_s)dz)$ is the rate at which individual i enters the job search assistance program after s periods. We use the delta method to compute standard errors around the treatment effects.

6 Estimation results

6.1 Average treatment effects on the treated survivors

In the discrete-time evaluation we fix the unit of time to 30 days. Table 2 presents estimated treatment effects using both the discrete-time and the continuous-time approach. In the table we consider the effect of entering the program in the second, fourth and eighth month of unemployment benefits on exit within three or eight months. This allows us to obtain insight in both the effects of entering the program early as well as late and on short-run and longer-run outcomes.

Both models estimate that participation in the job search assistance program does not stimulate the exit rate from unemployment. The discrete-time model shows negative effects of participating in the job search program on exit from the benefits scheme. The effects are significant and large when entering the program early. Individuals who enter the program during the second month of unemployment are about 10 percentage points less likely to exit unemployment between the beginning of the third month and the end of the fifth month. And the probability of leaving unemployment between the third and ninth month is 16.6 percentage points lower. The program effects are smaller and not significant when starting in the eight month of unemployment.

The discrete-time method requires specifying the unit of time. Long time intervals imply aggregating many transitions which, therefore, overlooks dynamics occurring within a time interval. However, choosing short time intervals reduces the number of observations leading to less precise estimates of transition probabilities. Furthermore, the interpretation of effects changes when treatment intervals are increased or decreased. Our choice

Table 2: Average treatment effects on the treated survivors in the full sample

$(\mathbf{s}, \mathbf{t})$ (in months)	$(\mathbf{2}, \mathbf{5})$	$(\mathbf{2}, \mathbf{10})$	$(\mathbf{4}, \mathbf{7})$	$(\mathbf{4}, \mathbf{12})$	$(\mathbf{7}, \mathbf{10})$	$(\mathbf{7}, \mathbf{15})$
Discrete-Time Logit						
$\mathrm{E}[Y^*_{0,t} \mid S = s, Y_s = 0]$	0.126	0.302	0.130	0.260	0.180	0.449
	(0.013)	(0.028)	(0.021)	(0.046)	(0.024)	(0.056)
$\mathrm{E}[Y^*_{1,t}(s) \mid S = s, Y_s = 0]$	0.025	0.137	0.059	0.243	0.170	0.403
	(0.011)	(0.027)	(0.017)	(0.043)	(0.026)	(0.034)
$\boldsymbol{\Delta}_{\mathrm{ATEST}}(\mathbf{t}, \mathbf{s})$	−0.100***	−0.166***	−0.071**	−0.017	−0.010	−0.046
	(0.017)	(0.036)	(0.028)	(0.062)	(0.034)	(0.061)
N	2411	2411	1588	1588	771	771
N (treated)	90	90	136	136	87	87
Continuous-time duration model						
$\mathrm{E}[Y^*_{0,t} \mid S = s, Y_s = 0]$	0.178	0.363	0.184	0.348	0.183	0.470
	(0.010)	(0.017)	(0.010)	(0.017)	(0.012)	(0.025)
$\mathrm{E}[Y^*_{1,t}(s) \mid S = s, Y_s = 0]$	0.123	0.267	0.126	0.252	0.124	0.349
	(0.010)	(0.017)	(0.010)	(0.016)	(0.010)	(0.018)
$\boldsymbol{\Delta}_{\mathrm{ATEST}}(\mathbf{t}, \mathbf{s})$	−0.055***	−0.096***	−0.058***	−0.096***	−0.059***	−0.121***
	(0.011)	(0.020)	(0.012)	(0.020)	(0.012)	(0.026)
N	3064	3064	3064	3064	3064	3064
N (treated)	788	788	788	788	788	788

Note: The unit of time in the discrete-time model is 30 days. Standard errors for the discrete-time model are obtained using 300 subsample replications, and for the continuous model using the Delta-method.
*** indicates significance at the 1% level, ** significance at the 5% level, * significance at the 10% level.
The discrete and continuous-time models include as covariates *gender*, *age*, *age*^{2} *log(pre-unemployment wage)*, *duration until exhaustion of benefits*, *I(unemployed in July or August)*, *I(age above 50)*, *I(low skilled)*.

for the size of a time interval $\kappa = 30$ was guided by the need for a large enough sample of treated individuals while limiting as much as possible dynamic selection within a single interval. In general, parameter estimates and standard errors are sensitive to the choice for the unit of time.

The continuous-time model finds significantly negative effects of participating in the job search model at all moments of entering and for all outcome lengths.[11] This is not surprising since the continuous-time model specifies the program effect as a homogeneous multiplicative effect on exit rates which means the effect of treatment is smoothed over all moments of entrance. The model still does capture that effects are smaller in magnitude for early entrance into the training program but larger for later entrance.

The results also show that the difference between both models in predicted potential treated outcomes $Y^*_{1,t}(s)$ is larger than that of the potential untreated outcomes. Part of this difference may be because the discrete time estimates do not impose a strong parametric assumption on the effect of the treatment and covariates on exit. The continuous-time

[11]Parameter estimates are presented in the appendix Table 5. To evaluate possible lock-in effects we also show parameter estimates for a model allowing for different effects depending on the time since beginning the training program. The effects on the hazard are however not significantly different from each other or from the overall effect.

model relies on proportional hazard rates, implying that the piecewise constant baseline hazard accounts for all changes in exit rates during the unemployment spell. If covariate effects on program entrance and job exit rates do not remain constant over the unemployment duration or are different between the treated and non-treated, then the continuous time model will not adequately capture time-varying heterogeneity.[12] The larger difference in treated outcomes may also be partly because the discrete time model estimates of the potential treated outcomes are based on a relatively small sample.

6.2 Sensitivity analysis

In this subsection we provide a number of sensitivity analyses. We mainly focus on the discrete-time model. Additional sensitivity analyses for the continuous-time model can be found in Kastoryano and Van der Klaauw (2011). In that working paper version we showed that there is some heterogeneity in treatment effects, mainly between low-skilled and regular worker and by elapsed unemployment duration when entering the program.

In Table 3 we compare our baseline discrete-time results to those from models with different control groups. The first frame reproduces our baseline discrete-time model. The second frame presents the results from Fredriksson and Johansson (2008). The control group in this method computes Kaplan-Meier estimates does not adjust estimates for simultaneous entrance into treatment in the non-treated survival probability. This will induce a dynamic selection bias if transitions into treatment and out of unemployment depend on an overlapping set of covariates as in our setting. The results indicate that omitting endogenous entrance into the job search program in the intervals $[\tau_{s+1}, \tau_{t-1}]$ increases the magnitude of the treatment effects.

In the third frame, we present results following Sianesi's (2004) definition of the control group which includes all individuals treated in the intervals $[\tau_{s+1}, \tau_{t-1}]$. Including these individuals will bias the estimates unless there is (on average) no effect of the treatment in this interval. As opposed to the two previous approaches, this approach does not cumulatively build up exit probabilities but estimates them over the entire period $[\tau_{s+1}, \tau_t]$. In general, we find increases in the treated exit probability. This is because the estimation is taken over the entire period $[\tau_{s+1}, \tau_t]$ which means censored observations are excluded and no information from those individuals is used when estimating exit probabilities. Since censored observations are more likely to be longer spells, ignoring these will produce an upward bias on the estimate of the treated exit probability.[13] This problem also influences the non-treated exit probability. However, the upward bias due to throwing out censored observations is counteracted by including the later treated in the control group whose

[12]We tried stratifying the duration dependence on different covariate structures. Estimating such models is complicated since the parameter space increases rapidly without any prior indication on how to stratify the duration dependence.

[13]This is also because in this method exit probabilities are not cumulatively built up so the missing at random assumption on censored observations is more likely to be violated.

Table 3: Comparing Matching Estimators for ATEST in the full sample

$(\mathbf{s},\mathbf{t})$ (in months)	$(\mathbf{2},\mathbf{5})$	$(\mathbf{2},\mathbf{10})$	$(\mathbf{4},\mathbf{7})$	$(\mathbf{4},\mathbf{12})$	$(\mathbf{7},\mathbf{10})$	$(\mathbf{7},\mathbf{15})$
Discrete-Time Logit (baseline)						
$E[Y^*_{0,t}\|S=s,Y_s=0]$	0.126	0.302	0.130	0.260	0.180	0.449
	(0.013)	(0.028)	(0.021)	(0.046)	(0.024)	(0.056)
$E[Y^*_{1,t}(s)\|S=s,Y_s=0]$	0.025	0.137	0.059	0.243	0.170	0.403
	(0.011)	(0.027)	(0.017)	(0.043)	(0.026)	(0.034)
$\boldsymbol{\Delta}_{\text{ATEST}}(\mathbf{t},\mathbf{s})$	−0.100***	−0.166***	−0.071**	−0.017	−0.010	−0.046
	(0.017)	(0.036)	(0.028)	(0.062)	(0.034)	(0.061)
N	2411	2411	1588	1588	771	771
N (treated)	90	90	136	136	87	87
No heterogeneity in exit probabilities						
$E[Y^*_{0,t}\|S=s,Y_s=0]$	0.135	0.341	0.147	0.288	0.192	0.475
	(0.015)	(0.033)	(0.025)	(0.050)	(0.029)	(0.057)
$E[Y^*_{1,t}(s)\|S=s,Y_s=0]$	0.025	0.137	0.059	0.243	0.170	0.403
	(0.011)	(0.027)	(0.017)	(0.043)	(0.026)	(0.034)
$\boldsymbol{\Delta}_{\text{ATEST}}(\mathbf{t},\mathbf{s})$	−0.110***	−0.204***	−0.088***	−0.044	−0.022	−0.072
	(0.019)	(0.040)	(0.031)	(0.065)	(0.035)	(0.062)
N	2411	2411	1588	1588	771	771
N (treated)	90	90	136	136	87	87
Including later entry in treatment						
$E[Y^*_{0,t}\|S=s,Y_s=0]$	0.133	0.334	0.125	0.281	0.201	0.468
	(0.015)	(0.027)	(0.018)	(0.033)	(0.027)	(0.040)
$E[Y^*_{1,t}(s)\|S=s,Y_s=0]$	0.026	0.145	0.059	0.327	0.171	0.440
	(0.011)	(0.028)	(0.018)	(0.050)	(0.027)	(0.037)
$\boldsymbol{\Delta}_{\text{ATEST}}(\mathbf{t},\mathbf{s})$	−0.108***	−0.189***	−0.066***	0.046	−0.030	−0.027
	(0.018)	(0.038)	(0.025)	(0.060)	(0.039)	(0.049)
N	2411	2411	1588	1588	771	771
N (treated)	90	90	136	136	87	87
Excluding later treated from control group						
$E[Y^*_{0,t}\|S=s,Y_s=0]$	0.143	0.606	0.174	0.515	0.253	0.602
	(0.016)	(0.056)	(0.030)	(0.074)	(0.038)	(0.072)
$E[Y^*_{1,t}(s)\|S=s,Y_s=0]$	0.026	0.145	0.059	0.327	0.172	0.446
	(0.011)	(0.028)	(0.018)	(0.046)	(0.027)	(0.037)
$\boldsymbol{\Delta}_{\text{ATEST}}(\mathbf{t},\mathbf{s})$	−0.117***	−0.460***	−0.115***	−0.188**	−0.081	−0.156**
	(0.019)	(0.062)	(0.034)	(0.091)	(0.046)	(0.076)
N	2245	1870	1447	1205	656	591
N (treated)	90	90	136	136	87	87

Note: The unit of time in the discrete-time model is 30 days. Standard errors for the discrete-time model are obtained using 300 subsample replications.
*** indicates significance at the 1% level, ** significance at the 5% level, * significance at the 10% level.
All models include include as covariates *gender*, *age*, *age*^{2} *log(pre-unemployment wage)*, *duration until exhaustion of benefits*, *I(unemployed in July or August)*, *I(age above 50)*, *I(low skilled)*.

treatment effect is negative. As a result, the non-treated exit probability is similar or smaller than in the second frame.

In the last frame we look at results when excluding from the control group unemployed workers who receive later treatment. As mentioned previously, excluding these individ-

Table 4: Heterogeneity in ATEST

$(\mathbf{s}, \mathbf{t})$ (in months)	$(\mathbf{2}, \mathbf{5})$	$(\mathbf{2}, \mathbf{10})$	$(\mathbf{4}, \mathbf{7})$	$(\mathbf{4}, \mathbf{12})$	$(\mathbf{7}, \mathbf{10})$	$(\mathbf{7}, \mathbf{15})$
Discrete-Time Logit: regular workers						
$\mathbf{\Delta}_{\text{ATEST}}(\mathbf{t}, \mathbf{s})$	-0.148^{***}	-0.201^{**}	-0.112^{***}	-0.023	-0.031	-0.016
	(0.032)	(0.074)	(0.039)	(0.071)	(0.040)	(0.051)
N	2050	2050	1325	1325	674	674
N (treated)	30	30	67	67	55	55
N (treated 50-)	7	7	3	3	38	38
N (treated 50+)	23	23	64	64	17	17
Continuous-time duration model: regular workers						
$\mathbf{\Delta}_{\text{ATEST}}(\mathbf{t}, \mathbf{s})$	-0.075^{***}	-0.120^{***}	-0.075^{***}	-0.118^{***}	-0.067^{***}	-0.140^{***}
	(0.015)	(0.024)	(0.015)	(0.024)	(0.014)	(0.030)
N	2663	2663	2663	2663	2663	2663
N (treated)	480	480	480	480	480	480
$\mathbf{\Delta}_{\text{ATEST}}(\mathbf{t}, \mathbf{s})\vert\mathbf{50}-$	-0.112^{***}	-0.152^{***}	-0.111^{***}	-0.153^{***}	-0.095^{***}	-0.170^{***}
	(0.027)	(0.040)	(0.027)	(0.039)	(0.024)	(0.045)
N	2092	2092	2092	2092	2092	2092
N (treated)	191	191	191	191	191	191
$\mathbf{\Delta}_{\text{ATEST}}(\mathbf{t}, \mathbf{s})\vert\mathbf{50}+$	-0.066^{***}	-0.114^{***}	-0.065^{***}	-0.112^{***}	-0.053^{***}	-0.122^{***}
	(0.021)	(0.038)	(0.022)	(0.039)	(0.020)	(0.045)
N	571	571	571	571	571	571
N (treated)	289	289	289	289	289	289

Note: The unit of time in the discrete-time model is 60 days. Standard errors for the discrete-time model are obtained using 300 subsample replications, and for the continuous model using the Delta-method.
*** indicates significance at the 1% level, ** significance at the 5% level, * significance at the 10% level.
All models include include as covariates *gender*, *age*, *age*^{2} *log(pre-unemployment wage)*, *duration until exhaustion of benefits*, *I(unemployed in July or August)*, and *I(age above 50)* for the two first frames.

uals will remove longer unemployment spells from the control group. This is evident in the results where we see un upward bias in the non-treated exit probabilities leading to treatment effects far larger in magnitude. This approach also does not generate exit probabilities cumulatively which results in the same upward bias in exit outcomes as in the third panel.

In Table 4 we consider heterogenous effects of the treatment for different subpopulations. The first two frames show discrete and continuous-time results including only regular workers, which implies excluding about 400 low-skilled workers from the total sample, many of whom were treated. Both the discrete-time and continuous-time results indicate that the job search program has worse effects on regular workers. These are mainly driven by the adverse effects for individuals entering the program early in their unemployment spell.

Recall also that the assignment rule differs between workers who were above and below age 50 when becoming unemployed. Most workers under age 50 do not enroll in the job search assistance program within the first six months of unemployment. We, therefore, provide a separate analysis for regular workers above and below age 50. Estimation in the discrete-time case is not possible due to the small sample of treated individuals. The continuous-time results on each group show that the treatment effect is stronger for individuals aged below 50. However, these differences may simply be due to a combination of two factors. First, that more adversely selected individuals will enter treatment in the below 50 group given the later assignment. Second, because the continuous-time model does not condition on survivors it will smooth effects of later entrants over time due to the multiplicativity of treatment on the hazard rate.

7 Conclusions

In this paper we used data from a unique institutional setting to evaluate the effectiveness of a job search assistance program for unemployed teachers. The setting allows us to convincingly impose a conditional independence assumption. We used this in the empirical analysis to compare two different methods for dynamic treatment evaluation. We used a discrete-time method based on the g-computation algorithm (Robins, 1997) and a continuous-time duration approach (Ridder, 1986). Both methods show negative effects of participation in the job search assistance program on the exit from unemployment.

Recall that in our application we focus on individuals in the primary education sector collecting unemployment insurance benefits. Unemployed workers from the primary school sector differ from other unemployed workers, for example, in composition and where they search for new employment. For this group of unemployed workers we find that participating in the job search assistance program does not stimulate exit from unemployment. However, the job search assistance program is a general program provided by commercial training agencies and many unemployed workers in the private sector also participate in this program. The poor performance might be the consequence of a mismatch between the program and workers in the primary education sector rather than the program being ineffective in general. For example, the program might press participants to search for work in the general labor market, while unemployed workers in the primary education sector mainly search for teaching jobs at primary schools. An alternative explanation for the poor performance might be the lump-sum costs of participating in this program. This creates lower incentives for the commercial agency to ensure that program participants find work than pay-for-performance schemes which are usually offered by benefits agencies.

As a consequence of the results discussed in this paper, the job search assistance program has been modified in a number of ways. First, after two months of unemployment there is now an introductory meeting in which individuals are informed about the program. The Participation Fund indicates that this reduces the resistance to participate in the program. In the new setup, individuals only enter the program after having collected benefits for eight months. This is later than in the previous setup and there is no difference anymore for individuals below and above age 50. The previous results indicate that the program effect is less negative if entry is later during the unemployment spell and expenditures also decreased because less individuals actually enter the program. Finally, the job search assistance program now has some voluntary elements, so individuals have some discretion to choose their degree of assistance.

Even though in our case we can justify making a conditional independence assumption, this is not sufficient to identify dynamic models for treatment evaluation. Both our approaches also require adding a no-anticipation assumption. We have some additional information on invitation letters for the job search assistance program, which allow us to provide evidence in favor of the no-anticipation assumption. Many studies on dynamic treatment effects also rely implicitly on a no-anticipation assumption, but this is often not explicitly discussed nor is evidence on the validity of the assumption presented. A second issue is that dynamic matching estimators require weighting of survivor probabilities to deal with subsequent program participation. Matching to make population similar at the moment of program enrollment is not sufficient.

By imposing proportionality of the hazard rate the continuous-time model makes a stronger parametric assumption than the discrete-time model. As a consequence of this parametrization, the continuous-time model allows estimation of treatment effects at moments or for groups where the program is not implemented. In our case the model also produces a treatment effects for young workers who enter the program quickly in their unemployment spell (which is not observed in the data). The discrete-time model is more flexible but results are somewhat sensitive to the choice of the unit of time. Furthermore, alternative choices from the literature on which individuals enter the control group produces biases on survival probabilities and treatment effects.

References

Abbring, J.H. (2003), "Dynamic econometric program evaluation", Discussion Paper 804, IZA Bonn.

Abbring, J.H. and J.J. Heckman (2007), "Econometric evaluation of social programs, part III: Distributional treatment effects, dynamic treatment effects and dynamic discrete choice, and general equilibrium policy evaluation", Chapter 72 in J.J. Heckman and E. Leamer, editors, *Handbook of Econometrics, Volume 6B*, 5145-5303, Elsevier Science, Amsterdam.

Abbring, J.H., and G.J. van den Berg, (2003), "The nonparametric identification of treatment effects in duration models", *Econometrica* 71(5): 1491-1517.

Black, D., J. Smith, M. Berger and B. Noel (2003), "Is the threat of reemployment services more effective than the services themselves? Evidence from random assignment in the UI system", *American Economic Review* 93(4): 1313-1327.

Card, D. and D.R. Hyslop (2005), "Estimating the effects of a time-limited earnings subsidy for welfare-leavers", *Econometrica* 73(6): 1723-1770.

Card, D., J. Kluve and A. Weber (2010), "Active labour market policy evaluations: a meta-analysis", *Economic Journal* 120(548): 452-477.

Crépon, B., M. Ferracci, G. Jolivet and G.J. van den Berg (2010), "Analyzing the Anticipation of Treatments Using Data on Notification Dates", Discussion Paper 5265, IZA Bonn.

Doornik, J.A. (2007), "Object-Oriented Matrix Programming Using Ox". *3rd ed. London: Timberlake Consultants Press and Oxford*: www.doornik.com.

Eberheim, C., J.C. Ham and R.J. LaLonde (1997), "The impact of being offered and receiving classroom training on the employment histories of disadvantaged women", *Review of Economics Studies* 64(4): 655-682.

Fredriksson, P. and P. Johansson (2008), "Dynamic treatment assignment: the consequences for evaluations using observational data",*Journal of Business and Economic and Statistics* 26(4): 435-445.

Gerfin M. and M. Lechner (2002), "A microeconometric evaluation of the active labour market policy in Switzerland", *Economic Journal* 112(482): 854-893.

Gill, R.D. and J.M. Robins (2001), "Causal inference for complex longitudinal data: the continuous case", *Annals of Statistics* 29(6): 1785-1811.

Ham, J.C. and R.L. LaLonde (1996), "The effect of sample selection and initial conditions in duration models: evidence from experimental data on training", *Econometrica* 64(1): 175-205.

Ham, J.C. and A. Rea Jr. (1987), "Unemployment insurance and male unemployment duration in Canada", *Journal of Labor Economics* 5(3): 325-353.

Heckman, J.J. and S. Navarro (2007), "Dynamic discrete choice and dynamic treatment effects", *Journal of Econometrics* 136(2): 341-396.

Heckman, J.J. and B. Singer (1984), "A method for minimizing the impact of distributional assumptions in econometric models for duration data", *Econometrica* 52(4): 271-320.

Imbens, G.W. and J.M. Wooldridge (2009), "Recent developments in the econometrics of program evaluation", *Journal of Economic Literature* 47(1): 5-86.

Kastoryano, S. and B. van der Klaauw (2011), "Dynamic evaluation of job search assistance", Discussion Paper 5424, IZA Bonn.

Lalive, R., J.C. van Ours and J. Zweimüller (2008), "The impact of active labour market programmes on the duration of unemployment in Switzerland", *Economic Journal* 118(525): 235-257.

LaLonde, R.J. (1986), "Evaluating the econometric evaluations of training programs with experimental data", *American Economic Review* 76(4): 604-620.

Lechner, M. and C. Wunsch (2009), "Are training programs more effective when unemployment is high?", *Journal of Labor Economics* 27(4): 653-692.

Lok, J., R. Gill, A. van der Vaart and J. Robins (2004), "Estimating the causal effect of a time-varying treatment on time-to-event using structural nested failure time models", *Statictica Neerlandica* 58(3): 271-295.

Lok, J,J., M.A. Hernán and J.M. Robins (2007), "Optimal start of treatment based on time-dependent covariates", JSM Proceedings 1149-1160.

OECD (2010), "Employment Outlook", OECD, Paris.

Politis, N., and J. Romano (1994), "Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions," *The Annals of Statistics*, 22, 2031-2050.

Ridder, G. (1986), "An event history approach to the evaluation of training, recruitment and employment programmes", *Journal of Applied Econometrics* 1(2), 109-126.

Robins, J.M. (1986), "A new approach to causal inference in mortality studies with sustained exposure periods – application to control of healthy worker survivor effect", *Mathematical Modelling* 7(9-12): 1393-1512.

Robins, J.M. (1997), "Causal inference from complex longitudinal data", in M. Berkane (eds.), *Latent Variable Modelling and Applications to Causality. Lecture Notes in Statistics (120)*, Springer-Verlag, 69-117, New York.

Sianesi, B. (2004), "An evaluation of the Swedish system of active labor market programs in the 1990's", *Review of Economics and Statistics* 86(1): 133-155.

Van den Berg, G.J. and B. van der Klaauw (2006), "Counseling and monitoring of unemployed workers: theory and evidence from a controlled randomized experiment", *International Economic Review* 47(3): 895-936.

Young, J.G., L.E. Cain, J.M. Robins, E.J. O'Reilly and M.A. Hernán (2011), "Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula", *Statistics in Biosciences* 3(1): 119-143.

A P-weighting G-computation formula

The g-computation formula in the text lends itself to estimation by matching, regression or stratification. If researchers favor estimation by propensity score weighting, it is straightforward to re-rewrite the g-computation formula in terms of a joint distribution over the propensity to treatment. To see this, we rewrite the potential treated outcome from the text as follows $k \geq s+1$,

$$
\begin{aligned}
&\mathrm{E}[Y^*_{1k}(s)|D_{s-1}=0, Y_s=0, X=x] \\
&= 1 - \prod_{j=s+1}^{k} \Pr(Y_j = 0|\bar{D}_{j-1}(s) = \bar{d}_{j-1}(s), Y_{j-1}=0, X=x) \\
&= 1 - \prod_{j=s+1}^{k} \Pr(Y_j = 0|D_{j-1}=1, \ldots, D_s=1, D_{s-1}=0, Y_{j-1}=0, X=x) \\
&= 1 - \frac{\prod_{j=s+1}^{k} \Pr(Y_j=0, \bar{D}_{j-1}(s) = \bar{d}_{j-1}(s)|Y_{j-1}=0, X=x)}{\Pr(\bar{D}_s(s) = \bar{d}_s(s)|Y_s=0, X=x)} \\
&= 1 - \frac{\prod_{j=s+1}^{k} \Pr(Y_j=0, \bar{D}_{j-1}(s) = \bar{d}_{j-1}(s)|Y_{j-1}=0, X=x)}{\Pr(D_s=1|D_{s-1}=0, Y_s=0, X=x)}
\end{aligned}
$$

where the equalities follow from Bayes' rule, the definition of $\bar{D}_{j-1}(s) = \bar{d}_{j-1}(s)$, and the fact that $D_s = 1$ is an absorbing state so for $j = s+1, \ldots, k$, $\Pr(Y_{j+1}=0, D_{j-1}=1|D_j = 1) = \Pr(Y_{j+1}=0|D_j=1)$ and $\Pr(D_j=1, D_s=1) = \Pr(D_s=1)$.

We can similarly derive a weighting G-computation formula for the counterfactual outcomes. Following the same reasoning as above we have for $k \geq s+1$,

$$
\begin{aligned}
&\mathrm{E}[Y^*_{0k}|D_{s-1}=0, Y_s=0, X=x] \\
&= 1 - \prod_{j=s+1}^{k} \Pr(Y_j=0|D_{j-1}=0, Y_{j-1}=0, X=x) \\
&= 1 - \frac{\prod_{j=s+1}^{k} \Pr(Y_j=0, \bar{D}_{j-1}(j-1) \neq \bar{d}_{j-1}(j-1)|D_{j-2}=0, Y_{j-1}=0, X=x)}{\prod_{j=s}^{k-1} \Pr(\bar{D}_j(j) \neq \bar{d}_j(j)|D_{j-1}=0, Y_j=0, X=x)} \\
&= 1 - \frac{\Pr(Y_{s+1}=0, D_s=0|D_{s-1}=0, Y_s=0, X=x)}{\Pr(D_s=0|D_{s-1}=0, Y_s=0, X=x)} \cdot \prod_{j=s+2}^{k} \Pr(Y_j=0|D_{j-1}=0, Y_{j-1}=0, X=x)
\end{aligned}
$$

where the equalities follow again from Bayes' rule, the definition of $\bar{D}_{j-1}(s) \neq \bar{d}_{j-1}(s)$, and the fact that $\Pr(D_{j-1}=0|D_j=0) = 1$. For $k = s+1$ the counterfactual outcome takes the simple form,

$$
\mathrm{E}[Y^*_{0k}|D_{s-1}=0, Y_s=0, X=x] = 1 - \frac{\Pr(Y_{s+1}=0, D_s=0|D_{s-1}=0, Y_s=0, X=x)}{\Pr(D_s=0|D_{s-1}=0, Y_s=0, X=x)}
$$

B Additional tables

Table 5: Continuous-time model covariate estimates (full sample).

	Baseline		Lock-in	
	θ	θ_s	θ	θ_s
Treatment				
δ	-0.447^{***}			
	(0.093)			
$\delta_{T_{LI} \leq 60}$			-0.532^{***}	
			(0.150)	
$\delta_{60 < T_{LI} \leq 240}$			-0.276^{**}	
			(0.108)	
$\delta_{240 < T_{LI}}$			-0.681^{***}	
			(0.155)	
Low-skill	-0.845^{***}	1.683^{***}	-0.818^{***}	1.683^{***}
	(0.096)	(0.081)	(0.096)	(0.082)
50+	0.089	0.518^{***}	0.107	0.483^{***}
	(0.134)	(0.131)	(0.134)	(0.131)
Female	-0.060	-0.113^{*}	-0.073	-0.141^{*}
	(0.069)	(0.075)	(0.069)	(0.075)
Age_{tran}	-1.032^{**}	5.551^{***}	-1.084^{**}	6.245^{***}
	(0.451)	(1.027)	(0.448)	(1.039)
Age^2_{tran}	0.957^{*}	-3.212^{***}	0.996^{*}	-3.677^{***}
	(0.572)	(0.830)	(0.570)	(0.842)
log(Wage)	-0.270^{***}	0.663^{***}	-0.290^{***}	0.664^{***}
	(0.062)	(0.116)	(0.061)	(0.116)
LengthUIben/365	-0.667^{***}	-0.057	-0.678^{***}	-0.057
	(0.037)	(0.038)	(0.037)	(0.039)
$\text{Layoff}_{Jul-Aug}$	0.323^{***}	-0.152^{**}	0.302^{***}	-0.140^{**}
	(0.059)	(0.068)	(0.058)	(0.068)
Duration Dep.				
λ_{0-60}	-4.077^{***}	-11.463^{***}	-4.013^{***}	-11.594^{***}
	(0.172)	(0.431)	(0.169)	(0.433)
λ_{60-120}	0.416^{***}	1.163^{***}	0.434^{***}	1.058^{***}
	(0.060)	(0.129)	(0.060)	(0.129)
$\lambda_{120-180}$	0.269^{***}	1.622^{***}	0.296^{***}	1.528^{***}
	(0.072)	(0.127)	(0.073)	(0.127)
$\lambda_{180-240}$	0.650^{***}	2.145^{***}	0.667^{***}	2.043^{***}
	(0.077)	(0.127)	(0.077)	(0.127)
$\lambda_{240-300}$	0.389^{***}	1.584^{***}	0.395^{***}	1.513^{***}
	(0.107)	(0.166)	(0.107)	(0.166)
$\lambda_{300-360}$	0.119	1.708^{***}	0.151	1.623^{***}
	(0.149)	(0.191)	(0.149)	(0.193)
$\lambda_{360-480}$	1.087^{***}	0.686^{***}	1.162^{***}	0.672^{**}
	(0.105)	(0.271)	(0.105)	(0.271)
$\lambda_{480-656}$	0.907^{***}	0.848^{***}	1.157^{***}	0.594
	(0.142)	(0.399)	(0.154)	(0.416)
Loglikelihood	-6.163		-6.158	
Observations	3064		3064	

*** indicates significance at the 1% level, ** significance at the 5% level, * significance at the 10% level.
Standard errors in parentheses. Age_{tran}=(Age−20)/40.
Individuals from Low Skilled jobs excluded in estimation.
Duration dependence terms in continuous-time model are deviations from λ_{0-60}.

Table 6: Continuous-time model covariate estimates for heterogenous effects (regular workers).

	Regular workers		Above 50		Under 50	
	θ	θ_s	θ	θ_s	θ	θ_s
Treatment						
δ	-0.502^{***}		-0.543^{***}		-0.515^{***}	
	(0.104)		(0.146)		(0.177)	
50+	0.153	0.559^{***}				
	(0.143)	(0.176)				
Female	-0.017	-0.138	0.008	0.102	-0.061	-0.251^{*}
	(0.072)	(0.095)	(0.081)	(0.173)	(0.158)	(0.135)
Age_{tran}	-0.939^{**}	6.141^{***}	-1.497^{***}	8.950^{***}	-4.060	-1.745
	(0.469)	(1.309)	(0.559)	(2.712)	(28.739)	(23.167)
Age^2_{tran}	0.827	-2.786^{***}	2.020^{***}	-6.199^{**}	0.564	1.907
	(0.605)	(1.024)	(0.740)	(2.681)	(16.697)	(13.146)
log(Wage)	-0.260^{***}	0.456^{***}	-0.238^{***}	0.274	-0.258	0.578^{***}
	(0.063)	(0.130)	(0.069)	(0.207)	(0.180)	(0.188)
LengthUIben/365	-0.682^{***}	-0.076^{*}	-0.824^{***}	-0.124	-0.439^{***}	-0.038
	(0.038)	(0.040)	(0.052)	(0.165)	(0.054)	(0.047)
$\text{Layoff}_{Jul-Aug}$	0.305^{***}	-0.248^{**}	0.269^{***}	-0.390^{**}	0.428^{***}	-0.152
	(0.062)	(0.091)	(0.068)	(0.158)	(0.143)	(0.130)
Duration Dep.						
λ_{0-60}	-4.120^{***}	-11.913^{***}	-3.989^{***}	-12.311^{***}	-1.979	-8.286
	(0.176)	(0.546)	(0.192)	(0.848)	(12.331)	(10.096)
λ_{60-120}	0.460^{***}	1.620^{***}	0.439^{***}	0.138	0.761^{***}	1.809^{***}
	(0.061)	(0.198)	(0.065)	(0.575)	(0.207)	(0.220)
$\lambda_{120-180}$	0.336^{***}	1.930^{***}	0.328^{***}	1.344^{***}	0.602^{**}	2.140^{***}
	(0.074)	(0.202)	(0.079)	(0.473)	(0.236)	(0.227)
$\lambda_{180-240}$	0.703^{***}	2.655^{***}	0.753^{***}	3.692^{***}	0.710^{***}	1.689^{***}
	(0.080)	(0.200)	(0.087)	(0.404)	(0.258)	(0.266)
$\lambda_{240-300}$	0.375^{**}	2.048^{***}	0.457^{***}	3.234^{***}	0.263	0.968^{***}
	(0.115)	(0.229)	(0.126)	(0.428)	(0.324)	(0.358)
$\lambda_{300-360}$	0.170	2.320^{***}	0.111	3.843^{***}	0.619^{*}	0.831^{*}
	(0.163)	(0.241)	(0.194)	(0.434)	(0.346)	(0.439)
$\lambda_{360-480}$	1.135^{***}	1.305^{***}	1.287^{***}	2.815^{***}	0.998^{***}	0.426
	(0.115)	(0.319)	(0.134)	(0.523)	(0.278)	(0.494)
$\lambda_{480-656}$	0.988^{***}	0.907^{*}	1.019^{***}	2.061^{***}	1.051^{***}	0.109
	(0.162)	(0.513)	(0.203)	(0.796)	(0.331)	(0.747)
Loglikelihood	-5.902		-5.711		-6.359	
Observations	2663		2092		571	

*** indicates significance at the 1% level, ** significance at the 5% level, * significance at the 10% level.
Standard errors in parentheses. Age_{tran}=(Age−20)/40.
Individuals from Low Skilled jobs excluded in estimation.
Duration dependence terms in continuous-time model are deviations from λ_{0-60}.

C Supplementary Materials

STATA-package for dynamic evaluation methods methods: contains STATA code to perform dynamic treatment effects estimation based on discrete time matching methods. The code includes options for matching based on the propensity score, nearest neighbor matching, and propensity score weighting. The file also contains the data set used in the article. (zipped file)

Ox-package for dynamic evaluation methods methods: contains Ox code to perform dynamic treatment effects estimation based on discrete time matching methods. The code includes options for matching based on the propensity score, nearest neighbor matching, and propensity score weighting. The file also contains the data set used in the article. (zipped file)