Dynamic Evaluation of Job Search Training

Stephen Kastoryano* Bas van der Klaauw†

September 20, 2010

Abstract

This paper evaluates job search training for unemployment insurance recipients. We use a unique administrative data set on individuals entering unemployment from the primary school sector. The dynamic assignment of the treatment allows us to assess the effects of the treatment on reemployment using different methods. In particular, we use timing-of-events, propensity score and regression discontinuity methods. We provide an extensive discussion of the identifying assumptions underlying the different methods with a particular focus on the issue that assignment to training is a dynamic process. We use the estimation results to investigate the targeting efficiency of the job search training.

*University of Amsterdam, VU University Amsterdam, Tinbergen Institute.

†VU University Amsterdam, Tinbergen Institute.

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

Thanks to Pierre Koning and seminar participants at the University of Tilburg, University of Zurich, Netherlands Bureau of Economic Policy Analyses, University of Essex, EALE/SOLE-meeting in London and the ESWM in Shanghai for useful comments.

1 Introduction

Since the 1990s, many countries offer job search training programs to stimulate the reemployment of unemployed workers. Policymakers often consider this to be a necessary requirement in a system with relatively generous unemployment benefits. In their recent survey, Card, Kluve and Weber (2009) stress that job search assistance programs often have relatively good short-run effects. Also in the Netherlands job search training if offered very frequently. In fact, the Netherlands is one of the 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 very limited. This is particularly true for job search training programs. In this paper, we focus on job search training for unemployed workers in the Dutch primary education sector. As outcome variable we consider the exit from unemployment, which is also the key variable of interest to policymakers. Focussing on, for example, on wages is less interesting, since the majority of the individuals returns to a job in primary school sector where wages are determined by collective bargaining and are almost a deterministic function of the individual's age.

In the Netherlands, the use of randomized social experiments in social insurance schemes is uncommon, i.e. Van den Berg and Van der Klaauw (2006) describe the most recent experiment conducted in 1998/99. Because, we have to deal with non-experimental data, the empirical evaluation of job search training is non-trivial. First, the evaluation of job search training suffers from the usual selection problem that participation might be related to (unobserved) characteristics. Second, job search training often does not start immediately after an individual enters unemployment, and the start of the training differs between individuals. Such dynamic assignment implies that when staying unemployed sufficiently long, all individuals will eventually enter job search training. This complicates the evaluation of job search training (see Abbring and Heckman, 2007; for a survey on dynamic treatment evaluation). It causes dynamic selection, which implies that some individuals have already left unemployment at the moment they should enter job search training. The construction of a comparable control group is thus complicated.

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 training can depend upon the moment of offering job search training. For example, lock-in effects may be more substantial for unemployed workers with relatively favorable labor market prospects (or those who are still relatively short-term unemployed). We investigate to what extent the timing of entering job search training affects the effectiveness of the program. Obviously, this is informative about the targeting efficiency of the program and such knowledge may improve the profiling of unemployed workers. In the current policy debate this issue becomes more important. Due to the financial crisis the Netherlands is facing substantial budget cuts, which will also affect the expenditures on active labor market programs.

In our empirical analyses, we use administrative data from a unique institutional environment in which the assignment to training is very clearly described and allows for different evaluation methods. The participation in job search training depends only on a limited set of observable characteristics and there are some clear discontinuities. We exploit this when estimating the effect of job search training using propensity score matching and regression discontinuity estimators. We compare the results of these estimators to assessing the program's effectiveness using the timing-of-events estimator. Because the institutional environment guarantees that the underlying assumptions of the different estimators are satisfied, we obtain valuable insights in the performance (in a real life setting) of some of the most popular cross-sectional estimators for policy evaluation used in microeconometric research.

In the US interventions during unemployment often start at a fixed moment, e.g. Black, Smith, Berger and Noel (2003) 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 entry into labor market programs is often varying between individuals. Other than in the Netherlands, this is, for example, 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), and Germany (Lechner and Wunsch, 2009). A relatively large literature attempts to fit such a dynamic setting into the standard potential outcome model (e.g. Gerfin and Lechner, 2002; and Sianesi, 2004). Sianesi (2004) discusses the complication of finding a suitable control group in case all individuals will eventually enter a program. Considering those individuals who are observed not to have received treatment implies conditioning on future outcomes. Fredriksson and Johansson (2008) argue that not accounting for dynamic selection may bias treatment evaluation estimators.

This paper fits well within the recently growing literature on dynamic treatment evaluation, surveyed by Abbring and Heckman (2007). The contribution is not only empirical, but by comparing the different methods we also intend to make some methodological contributions. We discuss the implementation of various estimators in a dynamic setting and show that all dynamic evaluation methods rely on assuming that individuals do not anticipate the exact start of treatment. Empirical studies using the timing-of-events methodology often explicitly justify this assumption (e.g. Van den Berg, Van der Klaauw and Van Ours, 2004), but this is ignored in studies using other evaluation estimators. Propensity score matching methods mainly focus on the conditional independence assumption, which is often justified from the richness of the data. Gerfin and Lechner (2002) and Sianesi (2004), for example, argues that information of past labor market outcomes and subjective assessments of labor market prospects justify the conditional independence 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 that the conditional independence assumption is valid. We compare the results from timing-of-events and propensity score matching estimation with regression-discontinuity estimation. The implementation of regression-discontinuity estimation in a dynamic setting is non-standard. We follow Abbring and Van den Berg (2005) when applying regression-discontinuity estimation and as far as we know our paper is the first which uses the regression-discontinuity estimator for duration models in an empirical application.

In the empirical application we mainly focus on the ex-post effects of job search training, which is the causal effect of actually entering the training program. Using the content and the goal of the training program we try to decompose this ex-post effect into a lock-in effect and an improvement in job search. In particular, we exploit that the training is most intensive during the first eight weeks. Most treatment evaluation estimators are not informative on the ex-ante effects, which are the effects of being enrolled in a benefits scheme which includes a job search training program. When applying regression-discontinuity we compare individuals who should enter the program quickly after becoming unemployed with those who enter later during the spell of unemployment. By comparing reemployment rates between both groups already before actually entering the program, we can get some idea about the threat effect of the program which might provide some insight in the size of possible ex-ante effects.

The paper is organized as follows. In Section 2 we provide details about the relevant unemployment benefits scheme, and the job search training. Section 3 presents the data. In Section 4 we provide a general framework for dynamic treatment evaluation. We discuss timing-of-events estimation in Section .... Section ... deals with propensity score methods and Section ... presents the results from regression discontinuity estimation. Finally, Section 8 concludes.

2 Institutional setting

2.1 Unemployment insurance for the primary education sector

Our data concern former employees of Dutch primary education who are entitled to collecting unemployment insurance benefits. Primary education institutions, as all public sector institutions, must bear their own unemployment insurance risk. However, since primary education institutions are relatively small, they were forced in 1996 to participate in a sector fund, which is the 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 each of 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, unemployment insurance benefits decrease 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 new 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 to their standing in the labor market (e.g. vacation, sickness, pregnancy, etc.). If the individuals fails to comply to these rules, a sanction can be applied which implies in a temporary reduction of the benefits level.

Over the last few years, 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 difference between workers in the primary education sector and the private sector. About 80% of the workers in primary education sector are women, and the average age is somewhat higher than in the private sector.

2.2 Job search training

Since July 2005, institutions in the public sector are also responsibility of reintegrating their former employees. This implies that the Participation Fund became responsible for financing and organizing reintegration activities for former employees of the primary education sector. 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 which can also include some vocational training. Second, a short job search training program focussing on networking skills in addition to job application training. Unemployed workers under age 60 are obliged to participate in reintegration activities if these are offered to them. Individuals who refuse to participate will be sanctioned with a substantial reduction of their benefits. Participation in the training 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 training program. Most individuals aim at finding new work again in the primary education sector, but about one-third of the observed exits is towards employment outside this sector.

Training is only provided to individuals who receive benefits for at least eight hours per week, and with an entitlement period exceeding three month. 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 upon an age criteria. Individuals above age 50 (at the first day of unemployment) and (low-skilled) individuals who were previously employed in a subsidized job, should enter the training program immediately after starting collecting benefits. Individuals under age 50 and who are not low-skilled, enter the training program only after six months of unemployment.

Only 8% of all programs offered are short job search training 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 programs offered are regular programs. These are offered by 11 private firms providing all the same program. Once invited for the regular program, the benefit recipients can choose the provider but 75% of the individuals accept the default. The remaining 25% almost always opt for the training 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 meeting depends upon the needs of the individual. The first weeks are often more intense, with two to three meetings per week with the training officers. The total amount spend in these meetings is about one full working day per week. After this period, the trainees usually visit the training center once a week or every other week for a few hours. During this later stage, trainees receive weekly assignments to be discussed in the weekly meetings. The general goal is that after two months of training the trainees should start making successful job applications. However, participation in the training does not lower the job search requirements. While in the training program, the unemployed workers have to comply to same minimum requirements on job applications as when not being in the program. And noncompliance to these requirements may cause that the benefit agency imposes a sanction. The training 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 training is 500 euro per individual entering the training. The cost of the regular training program is 4000 euro for individuals above age 50 and low-skilled individuals, and 3750 euro for individuals below age 50.

The Participation Fund does not assign benefit recipients directly to training programs, but outsources this task to a separate firm. This firm never has any personal contact with the 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 for benefits, number of weekly hours of collecting benefits, and an indicator code for the previous employer.[1] Two weeks prior to the start of the training the individual receives a letter explaining that they should enroll into a program. This letter also offers individuals to select one of the

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

11 private providers of the programs.

In practice the policy guidelines concerning the timing of entering training were not followed very strictly. This was due to administrative and communication issues between the Participation Fund and the external firm.[2] 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 treatment. And, since the external firm never had any contact with benefit recipients, the variation in treatment assignment should be exogenous conditional on observed 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. Furthermore, individuals must have claimed benefits for more than eight hours per week in order to be obliged to participate in a job search training program. We also exclude individuals above age 60 since for them participation in the job search training program is voluntary.

From the data we drop three individuals who very often entered and exited unemployment during the observation period. We excluded 43 observation for which the date of entering the job search training program was unknown or prior to becoming unemployed. The latter might occur if the individual was still in the program from an earlier unemployment spell. Finally, we exclude 37 observation with an hourly wage in the previous job below three euro, 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 in much more spread over the year, although there is

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

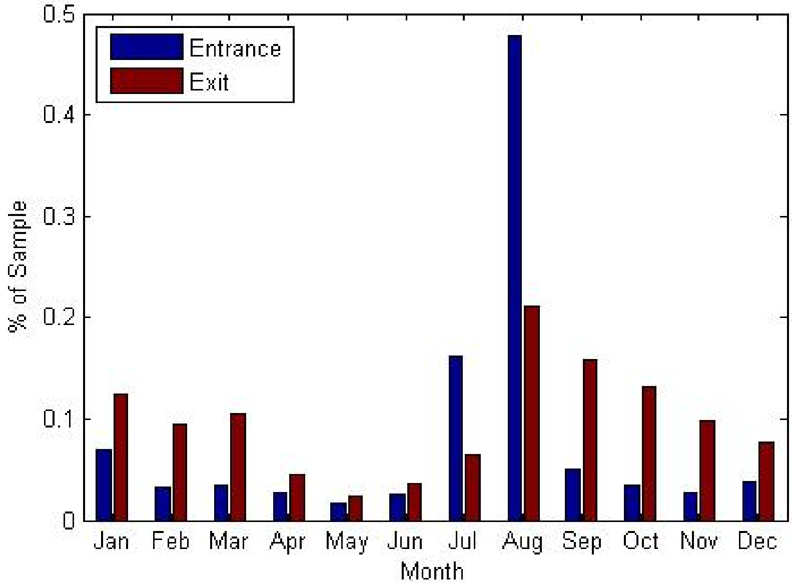
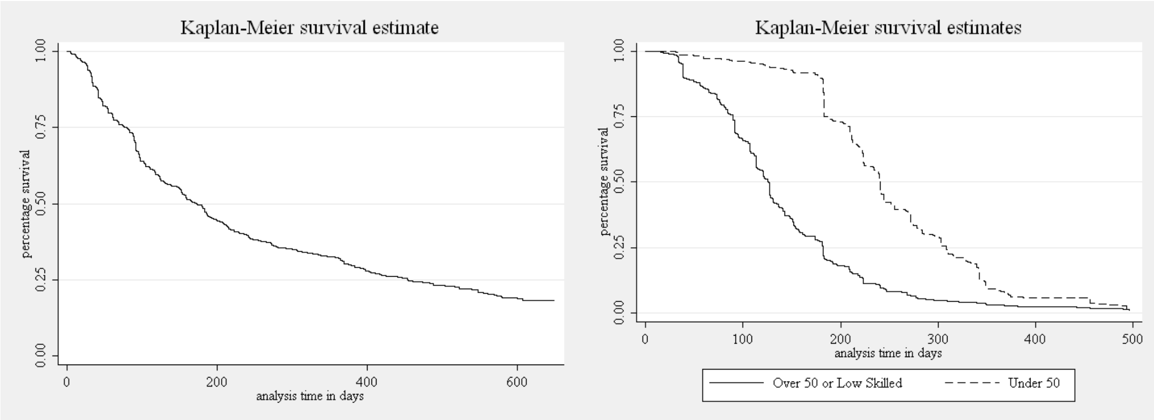


Figure 1: Seasonal variation in Entrance and Exit to unemployment

a decreasing trend over the school year. Figure 2a provides a Kaplan-Meier estimate for the exit to work. The median unemployment duration at about 21 weeks. Of the 3064 individuals, 862 entered the regular 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 training after six months of unemployment (below 50 and not low-skilled). The figure clearly shows that the latter group enters training, on average, later during the unemployment spell. Nevertheless, within each group there is still substantial variation in the moment of entering a program. 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 who assigned the programs. The nonparticipants are, on average, unemployed for more hours per week, and have a higher benefits level. This might be the direct consequence of the fact that many low-skilled individuals enter a program. Furthermore, also older workers are more likely to receive training, which follows the policy of assigning the programs. Of course, the different composition between the participants and the nonparticipants in not only the result of the assignment mechanism and the implementation of the external firm. Also dynamic selection plays an important role. Those individuals with adverse characteristics have, on



(a) K-M Estimate for Exiting Unemployment (b) K-M Estimate for Entering Training

Figure 2: Survival Estimates

Table 1: Descriptive statistics.

	Training participants	**Non-participants**
Number of observations	940	2124
Median unemployment duration (in days)	369	96
Median duration to training (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%

average, longer unemployment durations and are thus more likely to have entered training at some stage.

4 Model for dynamic treatment evaluation

4.1 Theoretical framework

In this section we briefly discuss a model for dynamic treatment evaluation. Our discussion fits within the more general discussion provided by Abbring and Heckman (2007). We highlight some issues relevant for actually estimating dynamic treatment effects. In the next sections we apply different estimation methods, and provide a comparison of the empirical results.

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

We focus on a situation in which individuals can receive a single treatment only once during the period of unemployment. All individuals in the data are eligible for entering training. However, the timing at which individuals receive training might differ. Let $S > 0$ denote the elapsed unemployment duration at the moment of entering treatment. Individuals might actually leave unemployment before the moment of entering training. Let D denote an indicator for actually observing treatment, $D = I(S < T)$.

Ideally, one should measure the effect of treatment at time period S on the residual unemployment duration $(T - S|T > S)$. This can easily be translated into a cost-benefit analysis in which the training costs are compared to future benefits payments. Furthermore, considering this treatment effect for different values of S is useful for improving the targeting of training to unemployed workers. However, data usually describe a limited time period, so long unemployment durations are censored. The lack of observations in the right tail of the distribution of unemployment spells implies that we cannot estimate average durations. Instead, we focus on whether or not someone is still unemployed some period after providing treatment.

Let $Y^*_{1,t}(s)$ denote the potential unemployment status after t periods if the individual was treated after s periods. So even though we only consider a single treatment, it may have different effects when applied at different time periods. We impose that treatment only affects later outcomes

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

There is thus no causal dependence of outcomes on future treatments. This is the no-anticipation assumption described by Abbring and Van den Berg (2003) and also adopted by Fredriksson and Johansson (2008). No-anticipation rules out that, prior to the actual start of job search training, unemployed workers already change their job search behavior in response to participating in the training. This rules out threat effects as, for example, measured by Black, Smith, Berger and Noel (2003). In their case unemployed workers their job search effort after having been informed about the actual start of a job search training program. Imposing no-anticipation does not rule out ex-ante effects of the treatment. Individuals may know that they are exposed to the risk of having to participate in some treatment, and may, therefore, behave differently than in a system in which the treatment is absent. Justifying the no-anticipation assumption requires knowledge about the unemployed worker's information about participation in job search training prior to actually starting the training. In our case, the unemployed workers are informed (by letter) two weeks prior the start of the training. Our data contain some information about these letters, which we exploit to justify the no-anticipation assumption (see Subsection 4.2).

In a dynamic setting it is not immediately clear what the relevant counterfactual is. The most natural counterfactual is to consider the potential outcome $Y^*_{0,t}$, which is the outcome if the unemployed would not receive treatment prior to t. This implies $S > t$, and because of the no-anticipation assumption

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

The relevant treatment effect (on the treated) would thus be

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

This treatment effect denotes the effect of providing treatment at s on reemployment between s and t for those who were still unemployed at s. This is the ex-post effect of the treatment, so the effect of actually participating in the treatment on future outcomes. It should be noted that almost all empirical literature focuses on ex-post effects. Unemployed workers treated at s are thus compared to unemployed workers who (possibly) receive training 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 identify which individuals did not receive treatment before t. In particular, for individuals who have left unemployment between s and t, it remains unobserved whether or not they 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. It may be clear that this causes a bias towards shorter unemployment spells, and treatment effects will be underestimated. Ignoring both types of observations does not solve the issue either as then there is no exit observed in the control group between s and t.

Sianesi (2004) suggests to consider as potential control group all individuals who receive treatment later than s. This implies that the treatment effect changes to

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

where $Y^*_{1,t}(> s)$ is the potential employment outcome at t for an unemployed worker treated later than s. This treatment effect describes the effect of entering treatment at s compared to 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 pattern of treatment. A cost-benefit analysis is, for example, not straightforward since it is unclear when individuals in the control group receive treatment. Both approaches to construct the 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. Abbring and Van den Berg (2003) use a duration model framework in which they jointly model the length of the unemployment spell T and the time until entry in treatment S. When imposing some functional form restrictions they can allow for selection on unobservables. Both Lechner (2008) and and Fredriksson and Johansson (2008) descretize time and develop multi-period matching estimators assuming that selection is only on observables. A more practical issue is the choice of the unit of a time interval. In the next sections of this paper we discuss in more detail the application of different approaches to our data, and we compare the estimation results. This should provide insight to the advantages and disadvantages of the different estimators.

4.2 Justifying no-anticipation

As stressed by Abbring and Van den Berg (2003), no-anticipation is required to identify the causal effects within the timing-of-events model. No-anticipation implies that individuals do not change their behavior prior to entering the program once they know the exact timing of starting the program. If unemployed workers receive information about the timing of job search training far before the actual start, they might take this into account in their current job search decisions. Indeed, Black et al. (2003) show that there may be substantial threat effect of active labor market programs, i.e. they find that unemployed workers are more likely to leave the benefits program once they are notified that they should start of a program.

Our data contain some information on invitation letters for the job search training. 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. Also there is 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 there is between 14 and 20 days between sending the letter and start of the job search training program.

The assumption of no-anticipation does not mean that individuals do not know about the assignment rules for job search training. Unemployed workers may be informed of the assignment rules to training. For example, an individual above age 50 may know that she should enter training as soon as possible, but it is ruled out that individuals know the exact timing of entering the training. This also implies that individuals cannot manipulate their assignment to job search training. Given the construction with the external agency assigning the job search training, it is unlikely that individuals can either manipulate or obtain prior knowledge about their assignment to the training. The 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 Timing-of-events approach

5.1 The model

We start by considering the timing-of-events approach proposed by Abbring and Van den Berg (2003) to estimate the effect of participating in the job search training program on reemployment. This is a continuous-time method which allows for selection on unobservables. The idea is to jointly model the reemployment rate T and the entry into treatment S in a bivariate mixed proportional hazard rate model. Bonnal, Fougère and Serandon (1997) use this model to estimate the effect of job search training, and Van den Berg, Van der Klaauw and Van Ours (2004) to evaluate the effectiveness of benefits sanctions.

Consider an individual collecting unemployment insurance benefits for t units of time. We assume that the exit rate to work can be characterized by observed characteristics x, unobserved characteristics v_u, the elapsed unemployment duration t itself, and a variable indicating whether the individual already started participating in job search training $I(s < t)$, where s is the moment at which an individual enters job search training. Furthermore, v_u is assumed to be independent of x. The transition rate from unemployment to work at t conditional on x, v_u and s is denoted by $\theta_u(t|x, v_u, s)$, and follows the familiar Mixed Proportional Hazard (MPH) specification

$$\theta_u(t|x, v_u, s) = \lambda_u(t) \exp(x'\beta_u + \delta \cdot I(s < t) + v_u)$$

in which $\lambda_u(t)$ represents the individual duration dependence.

The parameter δ describes the causal effect or participating in job search training on the exit rate from unemployment. The parameter δ describes a multiplicative effect on the reemployment rate. In Subsection 5.3 we show how we use this model to compute the dynamic treatment effects $\Delta(t, s)$ described in the previous section. Furthermore, in the model specification above δ is a permanent effect, which is the same for all individuals. This is most likely too strong. In Subsection 5.3 we allow δ to depend on individual characteristics x, the moment of starting job search training s, and the elapsed duration of job search training $t - s$.[3]

Recall that our data contain the same information as the external agency had when assigning unemployed workers to job search training. If we include all individual characteristics known to this external agency in the vector x, then the moment of entering the training s should be independent of unobservables v_u. We can test this by jointly modeling entry in the job search assistance and exit to work. Therefore, consider an individual who has received unemployment insurance benefits for t periods, and who did not start the job search training yet. The entry rate into job search training at t conditional on observed and unobserved characteristics x and v_s is denoted by

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

where x is again assumed independent of v_s. In this model specification, the entry rate in job search training is independent of the unobservables v_u only if v_s and v_u are independent. When actually estimating the model we allow for dependence between v_u and v_s via the joint distribution $G(v_u, v_s)$, and test for independence.

The identification of the model framework is discussed at length in Abbring and Van den Berg (2003). The identification highes on two key elements. First, proportionality in the hazard rates is necessary to identify the joint distribution of unobservables. This identification requirement thus imposes a restriction on the parametric specification. Second, the no-anticipation assumption discussed in the previous section is necessary. The model imposes a change in the exit rate from unemployment at the start of job search training. This implies that individuals should not already change their job search behavior prior to entering job search training in response to learning about the exact moment of entering the program.

The model is estimated using maximum likelihood estimation. Therefore, we parameterize the duration dependence functions and the bivariate unobserved heterogeneity distribution using flexible specifications. We take both $\lambda_u(t)$ and $\lambda_s(t)$ to have a piecewise constant specification,

$$\lambda_i(t) = \exp\left(\sum_{j=1,2,\ldots} \lambda_{ij} I_j(t)\right) \qquad i = u, s$$

where j is a subscript for time intervals and $I_j(t)$ are time-varying dummy variables for each of the consecutive time intervals. Note that with an increasing number

[3]Abbring and Van den Berg (2003) show that it is also possible to allow δ to depend on unobserved characteristics v.

of time intervals any duration dependence pattern can be approximated arbitrarily closely.

We take the joint distribution of the unobserved heterogeneity terms v_u and v_s to be bivariate discrete with unrestricted mass-point locations for each term. Allowing v_u and v_s to have two points of support each (v_u^a, v_u^b, v_s^a and v_s^b respectively), the distribution of $G(v_u, v_s)$ is

$$\begin{array}{ll} \Pr(v_u = v_u^a, v_s = v_s^a) = p_1 & \Pr(v_u = v_u^b, v_s = v_s^a) = p_3 \\ \Pr(v_u = v_u^a, v_s = v_s^b) = p_2 & \Pr(v_u = v_u^b, v_s = v_s^b) = p_4 \end{array}$$

with $0 \leq p_i \leq 1$ for $i = 1, \ldots, 4$, and $p_4 = 1 - p_1 - p_2 - p_3$. In this case it is easy to show that v_u and v_s are independent if and only if $\text{cov}(v_u, v_s) = 0$. The covariance of v_u and v_s equals

$$\text{cov}(v_u, v_s) = (p_1 p_4 - p_2 p_3) \cdot (v_u^a - v_u^b) \cdot (v_s^a - v_s^b)$$

In this case a zero covariance also implies independence between v_u and v_s, and thus conditional independence between assignment to job search training and exit to work. Of course, we can allow for more than two mass-point locations.

5.2 Parameter estimates

Table 2 presents the parameter estimates for the timing-of-events model, both for the full sample and for a sample excluding the low-skilled workers. In both cases the effect of participating in job search training is negative and significant. Entering a job search training program thus reduces the probability of finding work. In the full sample the reemployment rate drops about 29% = ($exp(-0.338) - 1$), and this is about 36% for regular unemployed workers (so excluding the low-skilled).

Low-skilled workers are less likely to exit unemployment to work and are significantly more likely to enter job search training. The latter is in agreement with the assignment policy of the program. Women are less likely to find work, but this is only significant in the full sample. The previous wage is negatively associated to finding work and positively to entering the job search assistance program. Furthermore, we allowed for a fourth-order polynomial in age. All terms have a significant effect both on reemployment and job search assistance participation.

Both in the exit rate to work and the entry in job search training, the duration dependence pattern is relatively flat beyond 30 days of collecting benefits. This implies that during the first month of unemployment not many people find new work, but later during the unemployment spell there is no decrease in the exit rate. The same holds for the entry in job search training.

The distribution of the unobserved heterogeneity is concentrated at a single mass point. So, there is no correlation between the unobserved heterogeneity v_u

in the reemployment rate and the unobserved heterogeneity v_s in the entry rate in job search training. This confirms that conditional of observed characteristics assignment to training is independent of the exit rate to work, which is in agreement with the process of assigning job search training as described in Subsection 2.2.

We have tried including additional heterogeneity in the model. In particular, we have tried including calender-time effects and regional dummies. However, most of the additional coefficients are insignificant and it did not change the parameter estimates. The latter is particularly true for the effects of participating in job search training.

Imposing proportionality of the hazard rates might be too strong. Low-skilled workers are very different from regular unemployed workers. Furthermore, individuals below and above age 50 have a different pattern on entry into training. Such differences are most likely not captured by a dummy variable causing a proportional shift in transition rates, and may imply very different duration dependence patterns. Therefore, we estimate the model separately for low-skilled workers, individuals above age 50 and those below age 50.

Table 3 shows the parameter estimates for the three groups. The job search training has a negative effect on reemployment for all three groups, but the effect is only significant for the no low-skilled individuals. Furthermore, the negative effect is almost twice as large for the individuals over age 50. Recall that individuals over age 50 are supposed to enter the job search training program early, while individuals below age 50 are supposed to wait about six month. Therefore, we interact the effect of the program with the elapsed unemployment duration when entering the program. Table 4 shows that for all groups job search training has the largest negative effect when the unemployed worker enters early, and that that this adverse effect becomes smaller the later the individuen enters the program.

It is well known that job search training programs may cause lock-in effects, i.e. when being in the program individuals reduce their job search effort. Recall that the objective of the job search training was to prepare someone for successful job search within two months and that the maximum length of the program is one year. Therefore, we allow the effect of the job search training to depend on the elapsed duration in the program and we allow the program effect to be different during the first two months, between two month and one year and beyond one year. Table 5 shows that for regular unemployed workers the effect of the program is negative at any elapsed duration since the start of the program. This indicates that the negative effects which we found earlier are not only the consequence of a very large initial lock-in effect.

5.3 Treatment effects

Within the timing-of-events model the causal effect of the job search training describes a shift in the exit rate to work. This does not directly translate into a statistic which is useful for policy purposes, such as the expected unemployment duration of employment probabilities. Since the timing-of-events model provides a full parametric specification for the exit rate to work and the entry into job search training, we could compute the effect of job search training on expected unemployment duration. However, this is unattractive since it requires extrapolating the pattern of duration dependence. Van den Berg and Van der Klaauw (2006) mention that expected unemployment durations are very sensitive on how the extrapolation is done.

Therefore, we focus on the treatment effects $\Delta(t, s)$ discussed in Section 4. These treatment effects describe the change in the probability of finding work within t periods of being unemployed due to entering job search training after s period (and conditional on still being unemployed at that moment). Using the timing-of-events model, this treatment effect can be written as

$$\Delta(t, s) = \frac{\exp(-\int_0^t \theta_u(z|x, t, v_u)dz) - \exp(-\int_0^t \theta_u(z|x, s, v_u)dz)}{\exp(-\int_0^s \theta_u(z|x, s, v_u)}$$

To obtain the population equivalents we average over the observed characteristics x of all individuals in our sample and integrate over the distribution of unobserved heterogeneity. We use the delta methods for computing standard errors around the treatment effects.

In Table we show the estimated employment probabilities and treatment effects. First we use the estimation results from the (baseline) timing-of-events model with permanent homogeneous effects of the job search training. The table first shows the effect of entering job search training after three months on employment after four and six months. The upper plane of the table shows that within three months about 30% of the individuals entering unemployment insurance find work. Without the job search training almost 39% of the individuals find work within four months and about 52 within six months. If individuals are assigned to job search training after three months, the employment probability after four months reduces to around 36% and almost 47% after six months. For $s = 3$ the treatment effects $\Delta(t, s)$ thus equal -0.037 and -0.081 for $t = 4$ and $t = 6$ respectively. Also for assignment to job search training after six months we find significantly negative effects on employment

We have also used the estimated models for the different groups. For low-skilled workers the effects are negative, but insignificant. For both other groups participation in job search training significantly reduces the probability of being employed. The treatment effects are about the same for individuals above and below age 50.

However, reemployment rates are much lower for individuals above age 50. Therefore, one might argue that individuals above age 50 suffer even more from participating in job search training.

Next, we considered the model in which the effect of the job search training was allow to depend on the elapsed unemployment duration at the start of job search training. The parameter estimates showed that the negative effect of job search training is largest for early entry in the program.

We have assumed that after the start of the training there is a continuous treatment effect. Recall from Subsection 2.2 that training is intense during the first weeks and that the goal is to prepare individuals for making successful applications after two months. There might be serious lock-in effects. E.g. Richardson and Van den Berg (2001) find negative overall treatment effect, but when ignoring the very strong lock-in effects positive effects of the treatment.

Table 6 shows the results for the separate groups taking treatment s at 3 months and 6 months with varying lengths for t. The results show that early treatment relative to the reemployment reference t increases with respect to the length of the period between s and t. We also notice that these differences are sharper at earlier stages of unemployment when many of the high ability workers are still in the sample of unemployed.

Figure 3 illustrates the evolution of the treatment effect in more detail for intervals [s,t] of 30, 90 and 180 days. Comparing the different series indicates that the choice of interval is not innocuous. We see a stronger effect for larger intervals which implies that differing assignment to treatment results in stronger reemployment prospects. The patterns also consistently show the reduction in the gains of postponing treatment until after t for individuals entering treatment at a later stage of their unemployment spell. But these patterns are surprisingly consistent for the different choices of intervals.

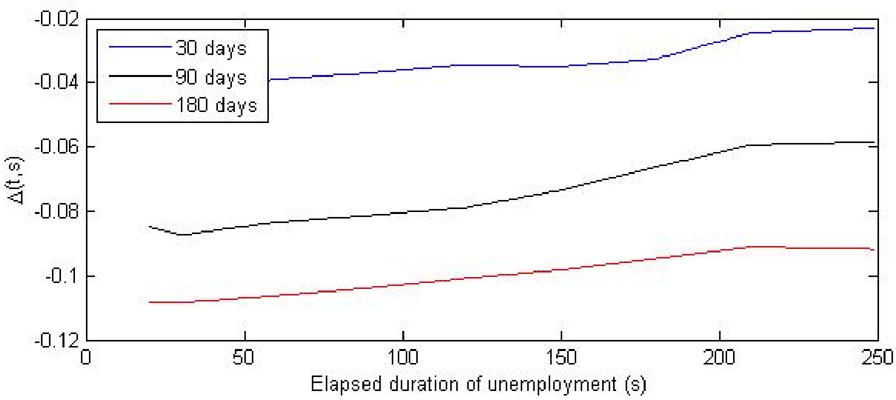


Figure 3: Timing of Events with different intervals [s,t]

6 Propensity score methods

The identification within the timing-of-events method relies on no-anticipation, and the assumption that hazard rates are proportional. The latter might be restrictive since it limits the parametric structure of the model. Recall that the timing-of-events model allows for selection on unobservables. From the institutional setting we know that in our situation, selection is only on observed characteristics. Conditional on observed characteristics, treatment assignment is thus independent of potential outcomes

$$(Y^*_{1,t}(s) \quad \forall t, s > 0) \perp S | X$$

Conditional independence assumption is usually justified from the richness of the data, which may include information on past labor market outcomes and subjective assessments of labor market prospects (e.g. Lechner and Gerfin, ...; and Sianesi, 2004).

Usual to Match treated at s with non-treated at s (i.e. Fredrikson and Johansson, 2008; Lalive, van Ours and Zweimüller, 2008; Gerfin and Lechner, ...; Sianesi, 2004). Should thus focus on individuals who survived the same period in unemployment and have similar observed characteristics. Overlapping support is thus required. Fredriksson and Johansson (2008) advocate the use of duration models in discrete time.

Conditional independence implies no-anticipation. Easy to see when considering that exit without treatment T_0 is a function of X and U and treatment assignment S of X and V with U and V independent. If anticipation than U also affects observed treatment assignment and conditional independence is violated.

......discretize time and assume that selection is only on observables. Lechner .... (....) suggest to simulate the training entry for those individuals for which the moment of treatment entry is missing (and then remove the treatment and control observations with a simulated treatment between s and t). Fredriksson and Johansson (2008) focus their attention on the effect of providing treatment at s compared to providing treatment at a later time period. This implies that their control group includes individuals who received treatment later than s rather than later than t. This complicates the interpretation of the treatment effect, since it does not take into account when control group individuals receive their treatment.

where $T^*(s)$ is the potential unemployment duration if job search training is assigned after s periods. With this independence assumption, we can use propensity score matching methods.

There are a few complicating issues. First, S describes the unemployment duration until entering job search training. And, since it is a continuous variable, there is an unlimited set of possible treatments to be evaluated. Imbens (2000) assesses this problem by estimating dose-response functions. But this framework applies indadequately to training program evaluations since for some treatments $S = s$, the potential outcome $T^*(s) < s$. This implies that if the treatment is assigned at s, the individual will have already left unemployment prior to actually starting the treatment. 74% of the spells in our data represent these individuals who exited unemployment prior to entering a training program.

Our propensity score mathcing estimator follows that of Lalive, Van Ours and Zweimüller (2008). Estimate survivor function both for the treatment and the matched control group. Focussing on survivor function solves the problem stressed by Sianesi (2004) that one can only compare current treatment start against later treatment start. This works fine due to the conditional independence assumption stressing that treatment start and exit from unemployment are independent processes conditional on observed characteristics.

To account for the conintuity of S, time is often discretized (e.g. Fredriksson and Johansson, 2008; Lechner, 2009; Sianesi, 2004). We normalize this unit to 1 day. In the empirical analyses we will discuss the sensitivity of the results with respect to choosing the unit of time. Treatment can then be written as $\tilde{S} = s$ if $s - 1 < S \leq s$ for $s = 1, 2, \ldots$. Recalling that $Y_t^*(s)$ denotes the potential employment status after t periods if the individual entered job search training after s periods, the conditional independence assumption in discrete time implies

$$\{Y_t^*(s) \forall s, t = 1, 2, \ldots\} \perp \tilde{S} | X$$

The large number of treatments, $\tilde{S} = 1, 2, \ldots$, is no longer a restrictive problem since we can now compare different treatments as

$$\mathrm{E}[Y_t^*(s) - Y_t^*(s') | \tilde{S} = s] \qquad \text{for} \quad s' \neq s$$

This gives the effect on reemployment at t for the unemployed workers who received job search training at s compared to receiving job search training at s'. Taking $s < s' < t$, one could typically think about this as the effect of being employed one year (t) after entering unemployment for those who enter job search training after three months (s) compared to six months (s'). The treatment statistic is thus informative about the effects of postponing the treatment from s to s'. On the other hand, by choosing $s < t \leq s'$, we can evaluate the effect of job search training at s on being employed at t compared to not having received treatment before t.

In order to make use of this treatment estimate we must account for the problem that we cannot identify both the group treated at s and at s'. This is due to the unobserved moment of entering treatment S for the individuals who left unemployment before entering job search training. We can partly solve this issue by conditioning the treatment evaluation on those people still collecting benefits at s,

$$\mathrm{E}[Y_t^*(s) - Y_t^*(s')|S = s, Y_s = 0] \qquad \text{for} \quad s' > s$$

Following this approach, two treatment parameters have been proposed. Sianesi (2004) argues that focus should lie on

$$\mathrm{E}[Y_t^*(s) - Y_t^*(> s)|\tilde{S} = s]$$

which is the effect of offering treatment at s compared to offering treatment at a later stage. This statistic is however not very useful for cost-benefit analysis or to assess the absolute performance of the job search training. It can only be relevant for caseworkers when deciding about the optimal timing of assigning treatment (conditional on actually assigning treatment).

Gerfin and Lechner (2002) take another approach and choose to drop individuals with observed treatment between s and t. Intuitively, it is clear that this causes an underestimate of the treatment effect. This is mainly because among those who should receive treatment between s and t, the short unemployment spells (found work before treatment) are kept while the long spells (treatment before finding work) are dropped.

Lalive, Van Ours and Zweimüller (2008) suggest a more suitable solution. Rather than directly focusing on whether or not the unemployed worker found work before t, they estimate survivor functions using a discrete-time duration model. Within this model, entering job search training between s and t is considered as exogenously right censored. This latter assumption is justified from the conditional independence assumption. The only requirement is that one conditions on the observed characteristics X which jointly affect entering treatment and exit to work.

In our study, we wish to compare the estimates produced by the timing of events methods to the matching ones proposed by Lechner (2009), Sianesi (2004), Fredriksson & Johansson (2008) and our own. In estimating the survival functions, and as opposed to Lechner (2009) and Sianesi (2004), we account for the imposed right censoring by further discretizing the interval between s and t. We then estimate the discrete time exit hazard for each subinterval and construct the survival probabilities. These exit hazards are first estimated as do Fredriksson & Johansson (2008) with Kaplan-Meier estimates and then, in our own application, with Logit specifications for each subinterval.

More specifically, for each period s, we compute the propensity score $Pr(S \in [s, s+\tau]|X, S \geq s, Y_s = 0)$ for all observations receiving treatment between $[s, s+\tau]$ (denoted $p(s)$), and the counterfactual group potentially receiving treatment at a time $s' > s + \tau$ (denoted $p(s')$). We then match each treated observation i to the 'nearest neighbour' j which minimizes $[p_i(s) - p_j(s')]^2$ in the control group.

In a second step, Fredriksson and Johansson (2008) compute the probability of exit for the treated group within each subinterval w_k of $[s,t]$, $w_k = [s, s+\tau), [s+\tau, s+2\cdot\tau), \ldots, [s+K\cdot\tau, t]$,

$$\widehat{H}_{1,w_k}(s) = \frac{\sum_{S\in[s,s+\tau]} Y_{1,w_k}(s)}{R_{1,w_k}(s)} = \frac{\sum_{S\in[s,s+\tau]} I(T\in w_k)}{\sum_{S\in[s,s+\tau]} I(T\geq w_k)}$$

and the potential unemployment status at t as,

$$Y^*_{1,t}(s) = 1 - \prod_{k=1}^{K}(1-\widehat{H}_{1,w_k}(s))$$

The counterfactual hazard and potential employment status using the matched control group are given by,

$$\widehat{H}_{0,w_k} = \frac{\sum_{S>s+\tau} Y_{0,w_k}}{R_{0,w_k}} = \frac{\sum_{S>s+\tau} I(T\in w_k)}{\sum_{S>s+\tau} I(T\geq w_k)}$$

and,

$$Y^*_{0,t} = 1 - \prod_{k=1}^{K}(1-\widehat{H}_{0,w_k})$$

In our estimation, $Y_{1,w_k}(s)$ and Y_{0,w_k}, are replaced by Logit functions in order to account for observed heterogeneity.

The final treatment effect is given by $\Delta(t,s) = Y^*_{1,t}(s) - Y^*_{0,t}$.

Lalive, van Ours and Zweimüller (2008) show that if one imposes conditional independence, propensity score matching and duration models give very similar results. However, when allowing for selection on unobservables the estimates effects change substantially. This might indicate that unobservables are important, even though their data are rich on individual characteristics, i.e. include past labor market outcomes and subjective assessments of labor market prospects. The latter are used by, for example, Lechner and Gerfin (...) and Sianesi (2004) to justify the conditional independence assumption.

6.1 Treatment Effects using Propensity Methods

One of the arguments for discretizing time and using matching methods rather than continuous time methods is that matching allows more precise interaction between durations and observables. In the duration model this interaction is introduced for $t > 0$ via the unobservables and the duration dependence terms. This specification may still fail to capture important variations over time.

In our study, we focus on the case when $s < t \leq s'$ and impose the right censoring approach for people entering training between s and t. Table 7 presents the results for matching taking $\tau = 30$ along with those of the different specifications of the timing of events model. To account for censoring in matching we also split the $t-s$ months hazards into K subintervals of length τ for the Fredriksson & Johansson and the Logit specifications. The choice for a matching interval and exit hazard subinterval may strongly influence the results. Our decision to apply the same interval τ is motivated by our willingness to decrease dynamic selection as much as possible subject to data limitations.

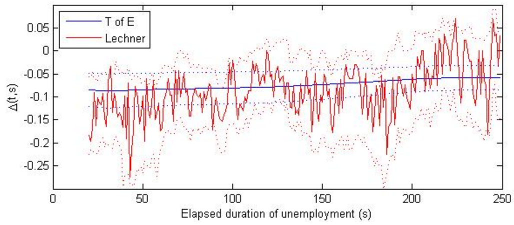
We notice that the main results from timing-of-events remain unchanged with only little variation between different approaches when taking s at three months. However, taking s at six months, and as the number of observations thins out, we notice that only the logit specification of the matching approach follows the results obtained in the timing of events. It is clear that the validity of the matching estimates depends importantly upon the matched sample of observations.

Figure 4 provides further insight into the different methods. We find that the matching methods produce a larger variation in the treatment effect. In particular, the estimates in 4a when dropping all observations with treatment between $(s+\tau, t]$ tend to underestimate the treatment effect. This follows from the bias selection of short spells which decreases the survival probability of the control group.

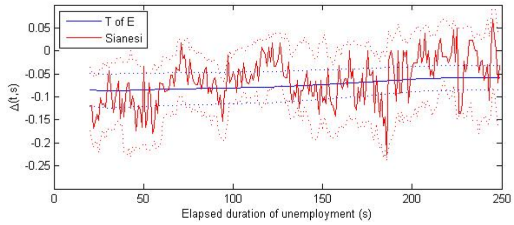
In the timing of events estimations, we noticed important heterogeneity based on observable characteristics. Kaplan-Meier estimates may not appropriately account for this observed heterogeneity in the survival probabilities. We can not therefore conclude that the fluctuations seen in 4b and 4c result solely from heterogeneity based on the determinants to treatment. In order to allow for the additional dependence of observables on exit, we replace the exit hazards in figure 4d by Logit estimates. Using the Logit specification, we find a strong reduction in the fluctuations of the treatment effect depending upon the elapsed duration of unemployment. These results nearly replicate those of the timing of events but still account for some additional duration dependent fluctuations based on observables.

7 Regression discontinuity

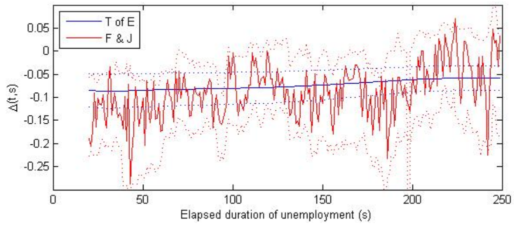
Regression discontinuity not straightforward in a dynamic setting with ongoing entry in job search training. Age affects the entry rate in training, which is known to individuals already at the start of unemployment. Forward-looking individuals use this information and, therefore, age may not be a valid instrument is our fuzzy setting. Rosenzweig and Wolpin (...) criticize the use of instrumental variables in dynamic settings. Abbring and Van den Berg (2005) arge that an indicator for being above or below the age threshold is a special regressor.



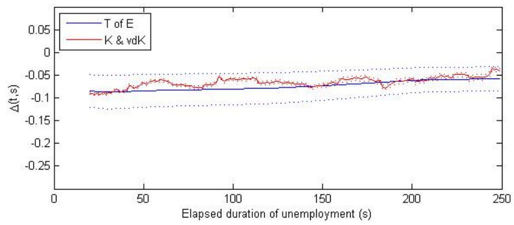
(a) Lechner



(b) Sianesi



(c) Fredriksson & Johansson



(d) Kastoryano & van der Klaauw

Figure 4: Treatment Effects for Propensity Matching Methods

Recall that the policy rule prescribes that individuals below age 50 and over age 50 at inflow into unemployment should be treated differently. Those individuals above age 50 were to enter job search training immediately after becoming unemployed, while individuals below age 50 were only to enter the program after six month of unemployment. However, the actual implementation of this policy failed. As presented in Section 3, there is substantial variation in the duration until entering job search training for each group, but on average individuals below age 50 enter job search training substantially later. We can exploit this partially failed policy using a fuzzy regression discontinuity design through instrumental variable methods.

The regression discontinuity presents a few complications when determining the causal effect of job search training. First, using instrumental variables in a dynamic setting with non-linear models is rather complicated. The complication is caused by the same issues as in the matching estimator, treatment assignment is an ongoing dynamic process and there is right-censoring. Also, individuals are likely to know when entering unemployment that they should participate in job search training fast (older than age 50 individuals) or after some period of unemployment (younger than age 50). Informed individuals might take this information into account when choosing their job search behavior, which may cause dynamic selection. This argument against the validity of instruments in dynamic settings is in line with Rosenzweig and Wolpin (JEL).

Abbring and Van den Berg (2005) propose a continuous-time and a discrete-time instrumental variable approach. The continuous-time approach closely follows the timing-of-events but with the exclusion restriction that being above or bellow 50 years old only affects the reemployment hazard up until the time of treatment. Taking z as the exclusion restriction with value $z = 0$ for individuals aged bellow 50, the MPH specification for the exit hazard becomes,

$$\theta_u(t|x, v_u, t_s) = \lambda_u(t) \exp(x'\beta_u + z'\eta_u \cdot I(t \leq t_s) + \delta \cdot I(t_s < t) + v_u)$$

Effect of the 50+ dummy.

This provides insight in the threat effects of having to enter the job search training early rather than later. Although these threat effects may provide some insight for the presence of ex-post effects, they should not be interpreted as ex-post effects. To measure ex-post effects, a group of individuals who will never enter the training is required. Note further that this threat effect differs from what, for example, Black et al. (2003) refer to as threat effect (which is in our context an anticipation effect).

The discrete time approach focuses on the survivor functions conditional upon the treatment status and the exclusion restriction. The exit probabilities in discrete time are given by,

$$Y^*_{0,t} = 1 - \frac{Pr(T > t, S = s'|z = 0) - Pr(T > t, S = s'|z = 1)}{Pr(S = s|z = 1) - Pr(S = s|z = 0)}$$

$$Y^*_{1,t}(s) = 1 - \frac{Pr(T > t, S = s|z = 1) - Pr(T > t, S = s|z = 0)}{Pr(S = s|z = 1) - Pr(S = s|z = 0)}$$

which identifies the dynamic MTE for the 'compliers' in the bellow age 50 group.

7.1 Treatment Effects using Regression Discontinuity Methods

Table 9 shows the results for the regression discontinuity in a dynamic framework. We limit the the sample to individuals aged between 40 and 60 and drop all individuals indicated as low skilled. The exclusion restriction in both the continuous-time and discrete-time approaches is an indicator for people older than 50 at the moment of layoff. In the discrete-time model, we focus on the results using the Logit specifications of the hazard calculated for subintervals over $\tau = 45$ days.

The results from the continuous-time regression discontinuity follow the timing of events results. We also see a larger relative difference in the treatment effect for longer intervals $[s, t]$. On the other hand, using a regression discontinuity approach in discrete time seems to largely diverge from previous results. It seems discretizing the data in a dynamic IV approach works poorly and the exclusion restriction performs quite badly.

Black et al. (2003) argue that the prospect of entering training may have strong effects on the unemployed workers job search behavior. Individuals knowing the future assignment date may increase their search efforts to avoid training. Figure 5 seems to give support to this argument. We see that the treatment effect is stronger for the marginal 'complier' group then in the timing of events or Logit matching models. This can be explained if the workers slightly younger than 50 years of age, and expecting to enter training 6 months into unemployment, exert less effort to find a job than their marginal counterparts.

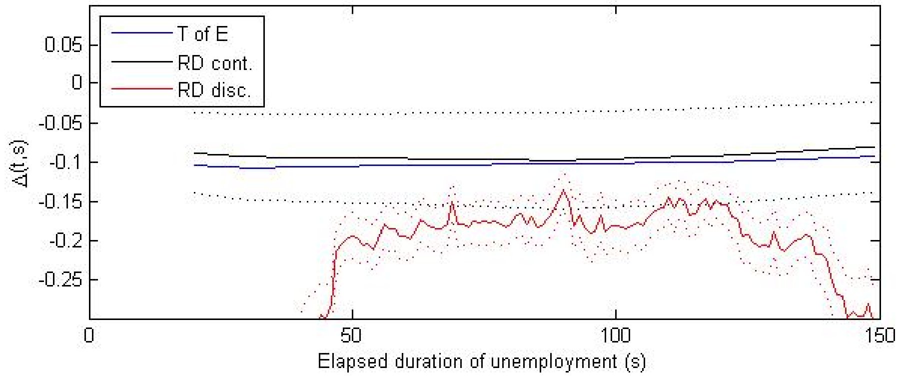


Figure 5: Comparing Treatment Effects across methods

8 Conclusions

Without imposing MPH identification of duration models with single spell data not possible. I.e. for each model with unobservables there exists an observationally equivalent without selection on observables, which allows for interactions between t and x.

Discrete-time methods have obvious disadvantages, i.e. choice of unit of time, continuous inflow, dealing with censoring, etc. Therefore, it might be much more difficult to fully capture the dynamic nature of the setting. Therefore, additional assumptions are required.

Key advantage of continuous-time methods is that no ad-hoc assumption on the unit of time should be made. More efficient estimators, since more data are used and easier to deal with issues as censoring and difference of entry in training. Disadvantage are parametric assumptions. However, latter seems modest as allowing for unobservables generates interaction between duration and observales.

Nonparametric propensity score methods and duration models may be observational relatively close. Propensity score methods do not allow for unobservables but impose less functional form assumptions and, therefore, such methods are relatively flexible on how duration dependence and observables interact. Mixed proportional hazard models do not allow directly for such interactions, but do allow for the presence of unobservables. The distribution of unobservables are identified from observed interactions between time and observables.

Card et al. (2009) shows that effectiveness of job search training is more favorable when evaluated using administrative data.

Card et al. (2009) stress that there is no bias in non-experimental studies compared to outcomes of experimental studies.

Unemployed workers in the primary school sector differ from the general population of unemployed workers. Different composition and much lower in ow rate. However, the aggregate exit rate from UI benefits is the same.

The job search assistance program evaluated in this paper is provided by private training companies. These companies offer the same programs also to UI benefits from the private sector.

The treatment effect is negative and significant, indicating that training reduces the exit rate to work. Several explanations can be given to this result. It may be that participating in training programs tends to give a negative signal to employers who in turn will be less likely to hire unemployed individuals following the training programs. Another possibility is that participating in training programs diminishes the amount of time available to search for a new position, this 'locking in' effect can be due to the length of the training programs or to the commuting time to the training centers. A last possibility is that the assignment of individuals to training simply causes them to decrease their search effort unconsciously or voluntarily by making people assume that the burden of searching for a new position is shared with the training program counselors.

estimate ex-post effect. ex-ante effect or threat effect remains unmeasured.

References

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., and G. Van den Berg, (2003), "The Nonparametric Identification of Treatment Effects in Duration Models", *Econometrica*, 71(5): 1491-1517.

Abbring, J.H. and G.J. Van den Berg (2004), "Analyzing the effect of dynamically assigned treatments using duration models, binary treatment models, and panel data models", *Empirical Economics*, 29, 5-20.

Abbring, J.H. and G.J. Van den Berg (2005), "Social experiments and instrumental variables with duration outcomes", Tinbergen Institute Discussion Paper 2005-047/3.

Bonnal, L., Fougre, D. and Serandon, A. (1997), "Evaluating the Impact of French Employment Policies on Individual Labour Market Histories", *Review of Economic Studies*, 64, 683-713.

Black, D., Smith, J., Berger, M., Noel, B., (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, 1313 1327.

Card, D., J. Kluve and A. Weber (2009), "Active Labor Market Policy Evaluations: a Meta-Analysis", *Working Paper*, no. 4002, IZA, Bonn.

Fredriksson, P. and P. Johansson (2008),"Dynamic Treatment Assignment-The Consequences for Evaluations Using Observational Data",*Journal of Business and Economic and Statistics*, 26, 435-445.

Gerfin M. and M. Lechner (2002), "A Microeconometric Evaluation of the Active Labour Market Policy in Switzerland", *Economic Journal*, 112, 854-893.

Hahn, J., P. Todd, and W. Van der Klaauw (2001), "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", *Econometrica* 69:1, 201-209.

Heckman, J., and S. Urzua. (2009), "Comparing IV With Structural Models: What Simple IV Can and Cannot Identify", *NBER Working Paper*, nbr. 14706.

Imbens, G. (2000). "The Role of the Propensity Score in Estimating Dose-Response Functions", *Biometrika* 87 (3): 706 - 710.

Lalive, R., Van Ours, J., Zweimller, J. (2008), “The impact of active labour market programmes on the duration of unemployment in Switzerland”, *Economic Journal* 118, 235-257.

Lechner, M, (2002), “Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies”, *Review of Economics and Statistics*, 84(2): 205-220, May.

Lechner, M., (2009), “Sequential Causal Models for the Evaluation of Labor Market Programs”, *Journal of Business and Economic Statistics*, 27, 71-83.

Lechner, M. and C. Wunsch (2009), ”Are Training Programs More Effective When Unemployment is High?”, *Journal of Labor Economics* 27(4): 653-692.

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

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., B. Van der Klaauw, and J.C. Van Ours (2004), “Punitive Sanctions and the Transition from Welfare to Work”, *Journal of Labor Economics*, 22, 211241.

Table 2: Timing of Events Estimates

	Full sample		No low-skilled	
	θ_u	θ_s	θ_u	θ_s
Treatment				
δ	-0.388^{**}		-0.443^{**}	
	(0.091)		(0.105)	
Gender	-0.115^{*}	-0.119	-0.061	-0.133
	(0.067)	(0.075)	(0.070)	(0.095)
Age/10	9.095^{**}	77.850^{**}	9.565^{**}	78.668^{**}
	(1.493)	(1.356)	(1.780)	(1.756)
$\text{Age}^2/10^2$	-4.293^{**}	-26.443^{**}	-4.590^{**}	-27.175^{**}
	(0.519)	(0.234)	(0.652)	(0.336)
$\text{Age}^3/10^3$	0.813^{**}	3.941^{**}	0.884^{**}	4.118^{**}
	(0.081)	(0.033)	(0.106)	(0.048)
$\text{Age}^4/10^4$	-0.055^{**}	-0.216^{**}	-0.061^{**}	-0.229^{**}
	(0.005)	(0.003)	(0.007)	(0.004)
Low Skilled	-0.918^{**}	1.618^{**}		
	(0.096)	(0.091)		
log(Wage)	-0.527^{**}	0.647^{**}	-0.507^{**}	0.498^{**}
	(0.055)	(0.116)	(0.057)	(0.129)
Duration Dep.				
λ_{30-90}	0.904^{**}	2.582^{**}	0.918^{**}	2.563^{**}
	(0.081)	(0.364)	(0.083)	(0.515)
λ_{90-150}	0.925^{**}	3.420^{**}	0.969^{**}	3.342^{**}
	(0.085)	(0.362)	(0.087)	(0.511)
$\lambda_{150-210}$	1.024^{**}	3.538^{**}	1.059^{**}	3.693^{**}
	(0.092)	(0.365)	(0.094)	(0.514)
$\lambda_{210-300}$	0.769^{**}	3.613^{**}	0.745^{**}	3.795^{**}
	(0.101)	(0.366)	(0.106)	(0.514)
$\lambda_{300-390}$	0.870^{**}	3.337^{**}	0.794^{**}	3.637^{**}
	(0.124)	(0.383)	(0.136)	(0.526)
$\lambda_{390-480}$	1.084^{**}	2.439^{**}	1.092^{**}	2.476^{**}
	(0.137)	(0.478)	(0.149)	(0.634)
$\lambda_{480-656}$	1.053^{**}	2.128^{**}	1.012^{**}	2.111^{**}
	(0.150)	(0.537)	(0.173)	(0.695)
Unobserved Het.				
v_1	-10.777^{**}	–	-11.048^{**}	–
	(1.560)		(1.780)	
v_2	–	-96.475^{**}	–	-96.020^{**}
		(2.430)		(3.064)
p_{11}	0.000		0.000	
	(–)		(–)	
p_{21}	0.000		0.000	
	(–)		(–)	
p_{12}	1.000		1.000	
	(–)		(–)	
p_{22}	0.000		0.000	
	(–)		(–)	
Loglikelihood	-18931.99		-15818.58	
Observations	3064		2663	

Note: ** indicates significance at the 1% level, * significance at the 5% level. Standard errors in parentheses

Table 3: Timing of Events Estimates by Policy Groups

	Only low-skilled		No low-skilled Over 50		No low-skilled Under 50	
	θ_u	θ_s	θ_u	θ_s	θ_u	θ_s
Treatment						
δ	-0.195		-0.875^{**}		-0.414^{**}	
	(0.285)		(0.219)		(0.148)	
Gender	-0.469^{*}	-0.113	-0.120	-0.131^{**}	-0.040	0.156
	(0.220)	(0.159)	(0.170)	(0.045)	(0.080)	(0.167)
Age/10	-24.256^{**}	-17.408^{**}	-49.799^{**}	7.563^{**}	32.359^{**}	-10.564
	(9.252)	(7.232)	(4.440)	(1.822)	(3.130)	(14.413)
$\text{Age}^2/10^2$	8.285^{**}	6.308^{**}	30.951^{**}	-73.093^{**}	-15.122^{**}	10.032^{*}
	(2.366)	(1.663)	(2.731)	(0.088)	(1.343)	(5.086)
$\text{Age}^3/10^3$	-1.291^{**}	-0.978^{**}	-5.848^{**}	17.472^{**}	2.993^{**}	-2.654^{**}
	(0.269)	(0.131)	(0.711)	(0.010)	(0.264)	(0.804)
$\text{Age}^4/10^4$	0.075^{**}	0.055^{**}	0.358^{**}	-1.183^{**}	-0.216^{**}	0.225^{**}
	(0.015)	(0.003)	(0.053)	(0.003)	(0.020)	(0.049)
log(Wage)	-1.408^{**}	0.687^{*}	-0.481^{**}	0.646^{**}	-0.507^{**}	0.382^{*}
	(0.332)	(0.377)	(0.197)	(0.077)	(0.062)	(0.202)
Duration Dep.						
λ_{30-90}	0.542	3.860^{**}	$0.632*$	2.456^{**}	0.945^{**}	0.000
	(0.380)	(0.631)	(0.290)	(0.518)	(0.086)	
λ_{90-150}	-0.142	5.444^{**}	1.159^{**}	3.460^{**}	0.949^{**}	0.815^{*}
	(0.440)	(0.645)	(0.306)	(0.514)	(0.091)	(0.453)
$\lambda_{150-210}$	0.282	5.266^{**}	0.912^{**}	3.571^{**}	1.100^{**}	2.861^{**}
	(0.454)	(0.650)	(0.386)	(0.521)	(0.099)	(0.358)
$\lambda_{210-300}$	0.664	5.217^{**}	0.779^{*}	3.301^{**}	0.784^{**}	3.434^{**}
	(0.453)	(0.661)	(0.454)	(0.535)	(0.112)	(0.341)
$\lambda_{300-390}$	0.958^{*}	3.799^{**}	1.014^{*}	3.383^{**}	0.806^{**}	3.578^{**}
	(0.470)	(0.859)	(0.521)	(0.590)	(0.152)	(0.367)
$\lambda_{390-480}$	0.837^{*}	4.291^{**}	1.302^{*}	2.446^{*}	1.129^{**}	2.669^{**}
	(0.492)	(0.883)	(0.569)	(1.117)	(0.171)	(0.551)
$\lambda_{480-656}$	0.976^{*}	3.965^{**}	1.382^{*}	5.779^{**}	0.997^{**}	1.817^{**}
	(0.483)	(0.984)	(0.666)	(0.509)	(0.209)	(0.762)
Unobserved Het.						
v_1	23.887^{*}	9.592	-22.978	–	-29.120^{**}	–
	(12.548)	(10.502)	(18.764)		(2.780)	
v_2	–	5.671	-30.912	335.300^{**}	–	-16.007
		(10.498)	(5865.900)	(6.521)		(15.192)
p_{11}	0.121^{**}		0.000		0.000	
	(0.030)		(–)		(–)	
p_{21}	0.000		0.000		0.000	
	(–)		(–)		(–)	
p_{12}	0.879^{**}		0.892		1.000	
	(0.030)		(0.916)		(–)	
p_{22}	0.000		0.108		0.000	
	(–)		(0.916)		(–)	
Loglikelihood	-3043.42		-3647.53		-12052.09	
Observations	401		571		2092	

Note: ** indicates significance at the 1% level, * significance at the 5% level.
Standard errors in parentheses.

Table 4: Heterogenous Treatment Effects

	Treatment Effect δ			
	Full Sample	*Low Skilled*	*Over 50 no LS*	*Under 50 no LS*
Inflow within 4 months	-0.678 **	-0.375	-1.194 *	-1.016 *
	(0.138)	(0.327)	(0.556)	(0.470)
Inflow 4 - 8 months	-0.257 **	0.039	-0.705	-0.425 **
	(0.108)	(0.306)	(0.546)	(0.176)
Inflow after 8 months	-0.166	0.037	-0.220	-0.171
	(0.175)	(0.526)	(0.615)	(0.219)

Note: ** indicates significance at the 1% level, * significance at the 5% level.
Standard errors in parentheses

Table 5: Locking-in Effects

	Treatment Effect δ			
	Full Sample	*Low Skilled*	*Over 50 no LS*	*Under 50 no LS*
$\delta_{t_{LI} \leq 60}$	-0.589 **	-0.406	-0.943 **	-0.540 *
	(0.151)	(0.383)	(0.274)	(0.234)
$\delta_{60 < t_{LI} \leq 365}$	-0.262 **	-0.086	-0.855 **	-0.286 *
	(0.102)	(0.310)	(0.174)	(0.163)
$\delta_{365 < t_{LI}}$	-0.568 **	0.014	-2.233 **	-0.972 *
	(0.221)	(0.458)	(0.426)	(0.562)

** indicates significance at the 1% level, * significance at the 5% level

Table 6: Heterogenous Treatment Effects by Policy Group.

	(s=3,t=4)	**(s=3,t=6)**	**(s=6,t=9)**	**(s=6,t=12)**
Full Sample				
$Y^*_{0,s}$	0.301 **	0.301 **	0.521 **	0.521 **
	(0.008)	(0.008)	(0.008)	(0.008)
$Y^*_{0,t}$	0.388 **	0.521 **	0.637 **	0.716 **
	(0.007)	(0.008)	(0.009)	(0.010)
$Y^*_{1,t}(s)$	0.362 **	0.465 **	0.606 **	0.670 **
	(0.009)	(0.013)	(0.009)	(0.010)
$\Delta(s,t)$	**-0.037 ****	**-0.081 ****	**-0.066 ****	**-0.096 ****
	(0.008)	**(0.018)**	**(0.015)**	**(0.022)**
Low Skilled				
$Y^*_{0,s}$	0.102 **	0.102 **	0.163 **	0.163 **
	(0.016)	(0.016)	(0.020)	(0.020)
$Y^*_{0,t}$	0.120 **	0.163 **	0.252 **	0.357 **
	(0.016)	(0.020)	(0.032)	(0.050)
$Y^*_{1,t}(s)$	0.117 **	0.153 **	0.237 **	0.328 **
	(0.016)	(0.019)	(0.020)	(0.023)
$\Delta(s,t)$	**-0.003**	**-0.011**	**-0.017**	**-0.035**
	(0.005)	**(0.016)**	**(0.027)**	**(0.053)**
Under 50 no LS				
$Y^*_{0,s}$	0.388 **	0.388 **	0.653 **	0.653 **
	(0.010)	(0.010)	(0.010)	(0.010)
$Y^*_{0,t}$	0.493 **	0.653 **	0.778 **	0.848 **
	(0.009)	(0.010)	(0.012)	(0.014)
$Y^*_{1,t}(s)$	0.460 **	0.583 **	0.743 **	0.803 **
	(0.013)	(0.024)	(0.016)	(0.019)
$\Delta(s,t)$	**-0.054 ****	**-0.114 ****	**-0.100 ****	**-0.130 ****
	(0.016)	**(0.037)**	**(0.032)**	**(0.034)**
Over 50 no LS				
$Y^*_{0,s}$	0.118 **	0.118 **	0.291 **	0.291 **
	(0.027)	(0.027)	(0.025)	(0.025)
$Y^*_{0,t}$	0.186 **	0.291 **	0.393 **	0.488 **
	(0.028)	(0.025)	(0.033)	(0.043)
$Y^*_{1,t}(s)$	0.147 **	0.197 **	0.336 **	0.385 **
	(0.027)	(0.026)	(0.024)	(0.029)
$\Delta(s,t)$	**-0.044 ****	**-0.107 ****	**-0.080 ****	**-0.145 ****
	(0.014)	**(0.025)**	**(0.024)**	**(0.038)**
Heterogeneity				
$Y^*_{0,s}$	0.302 **	0.302 **	0.523 **	0.523 **
	(0.008)	(0.008)	(0.008)	(0.008)
$Y^*_{0,t}$	0.389 **	0.523 **	0.640 **	0.716 **
	(0.007)	(0.008)	(0.009)	(0.011)
$Y^*_{1,t}(s)$	0.348 **	0.431 **	0.618 **	0.685 **
	(0.009)	(0.016)	(0.010)	(0.012)
$\Delta(s,t)$	**-0.058 ****	**-0.132 ****	**-0.046 ***	**-0.065 ***
	(0.009)	**(0.023)**	**(0.019)**	**(0.027)**
Locking-In				
$Y^*_{0,s}$	0.301 **	0.301 **	0.522 **	0.522 **
	(0.008)	(0.008)	(0.008)	(0.008)
$Y^*_{0,t}$	0.388 **	0.522 **	0.637 **	0.713 **
	(0.007)	(0.008)	(0.009)	(0.010)
$Y^*_{1,t}(s)$	0.352 **	0.456 **	0.599 **	0.670 **
	(0.010)	(0.014)	(0.009)	(0.010)
$\Delta(s,t)$	**-0.052 ****	**-0.095 ****	**-0.079 ****	**-0.089 ****
	(0.011)	**(0.019)**	**(0.016)**	**(0.022)**

** indicates significance at the 1% level, * significance at the 5% level

Table 7: Comparing Treatment Effects for Propensity Methods

Full Sample	**(s=3,t=4)**	**(s=3,t=6)**	**(s=6,t=9)**	**(s=6,t=12)**
Duration				
$Y^*_{0,s}$	0.301 **	0.301 **	0.521 **	0.521 **
	(0.008)	(0.008)	(0.008)	(0.008)
$Y^*_{0,t}$	0.388 **	0.521 **	0.637 **	0.716 **
	(0.007)	(0.008)	(0.009)	(0.010)
$Y^*_{1,t}(s)$	0.362 **	0.465 **	0.606 **	0.670 **
	(0.009)	(0.013)	(0.009)	(0.010)
$\Delta(s,t)$	**-0.037 ****	**-0.081 ****	**-0.066 ****	**-0.096 ****
	(0.008)	**(0.018)**	**(0.015)**	**(0.022)**
Lechner				
$Y^*_{0,s}$	0.307 **	0.307 **	0.563 **	0.563 **
	(0.002)	(0.002)	(0.002)	(0.002)
$Y^*_{0,t}$	0.332 **	0.427 **	0.653 **	0.739 **
	(0.016)	(0.022)	(0.021)	(0.025)
$Y^*_{1,t}(s)$	0.322 **	0.342 **	0.620 **	0.724 **
	(0.003)	(0.005)	(0.005)	(0.007)
$\Delta(s,t)$	**-0.014**	**-0.122 ****	**-0.076**	**-0.033**
	(0.024)	**(0.032)**	**(0.049)**	**(0.056)**
Sianesi				
$Y^*_{0,s}$	0.307 **	0.307 **	0.563 **	0.563 **
	(0.002)	(0.002)	(0.002)	(0.002)
$Y^*_{0,t}$	0.332 **	0.382 **	0.643 **	0.686 **
	(0.016)	(0.022)	(0.019)	(0.022)
$Y^*_{1,t}(s)$	0.322 **	0.342 **	0.620 **	0.724 **
	(0.003)	(0.005)	(0.005)	(0.007)
$\Delta(s,t)$	**-0.014**	**-0.058**	**-0.054**	**0.087**
	(0.024)	**(0.032)**	**(0.047)**	**(0.052)**
F & J				
$Y^*_{0,s}$	0.307 **	0.307 **	0.563 **	0.563 **
	(0.002)	(0.002)	(0.002)	(0.002)
$Y^*_{0,t}$	0.332 **	0.430 **	0.655 **	0.786 **
	(0.016)	(0.022)	(0.021)	(0.028)
$Y^*_{1,t}(s)$	0.322 **	0.348 **	0.620 **	0.737 **
	(0.003)	(0.005)	(0.005)	(0.007)
$\Delta(s,t)$	**-0.014**	**-0.117 ***	**-0.080**	**-0.113**
	(0.024)	**(0.033)**	**(0.050)**	**(0.064)**
K & VdK Logit				
$Y^*_{0,s}$	0.307 **	0.307 **	0.563 **	0.563 **
	(0.002)	(0.002)	(0.002)	(0.002)
$Y^*_{0,t}$	0.340 **	0.361 **	0.599 **	0.619 **
	(0.002)	(0.002)	(0.002)	(0.003)
$Y^*_{1,t}(s)$	0.308 **	0.314 **	0.571 **	0.588 **
	(0.002)	(0.002)	(0.003)	(0.003)
$\Delta(s,t)$	**-0.045 ****	**-0.068 ****	**-0.064 ****	**-0.071 ****
	(0.001)	**(0.002)**	**(0.003)**	**(0.004)**

Matching taken over 30 day interval [s,s+30)

** indicates significance at the 1% level, * significance at the 5% level

We use a quadratic specification for the age function in matching estimations

Table 8: Continuous Time IV Estimates

	Age between 40-60	
	θ_e	θ_p
Treatment		
δ	-0.591 **	
	(0.248)	
Gender	-0.233 *	-0.150
	(0.124)	(0.139)
Age/10	2.691	0.130
	(1.887)	(2.691)
$Age^2/10^2$	-0.352 *	0.050
	(0.193)	(0.256)
log(Wage)	-0.638 **	0.778 **
	(0.129)	(0.175)
Above 50	0.156	1.703 **
	(0.168)	(0.263)
Duration Dep.		
λ_{30-90}	0.740 **	2.542 **
	(0.171)	(0.521)
λ_{90-150}	1.089 **	3.658 **
	(0.180)	(0.518)
$\lambda_{150-210}$	1.151 **	4.438 **
	(0.224)	(0.527)
$\lambda_{210-300}$	0.833 **	5.637 **
	(0.272)	(0.547)
$\lambda_{300-390}$	1.030 **	6.040 **
	(0.327)	(0.593)
$\lambda_{390-480}$	1.440 **	4.935 **
	(0.346)	(1.011)
$\lambda_{480-656}$	1.371 **	5.685 **
	(0.389)	(0.885)
Unobserved Het.		
v_1	-9.407 *	-13.965 *
	(4.588)	(7.013)
v_2	-11.274 **	-18.349 **
	(4.603)	(7.024)
p_{11}	0.457 **	
	(0.108)	
p_{21}	0.197 *	
	(0.102)	
p_{12}	0.214 **	
	(0.077)	
p_{22}	0.132 *	
	(0.065)	
Loglikelihood	-6973.483	
Observations	1144	

** indicates significance at the 1% level
* significance at the 5% level
standard errors given in parenthesis

Table 9: Comparing Treatment Effects for Regression Discontinuity Methods

Age between 40-60	(s=3,t=4)	(s=3,t=6)	(s=6,t=9)	(s=6,t=12)
Cont.-time IV 40-60				
$Y^*_{0,s}$	0.179 **	0.179 **	0.379 **	0.379 **
	(0.011)	(0.011)	(0.015)	(0.015)
$Y^*_{0,t}$	0.257 **	0.379 **	0.486 **	0.567 **
	(0.012)	(0.015)	(0.019)	(0.024)
$Y^*_{1,t}(s)$	0.222 **	0.298 **	0.444 **	0.498 **
	(0.014)	(0.023)	(0.015)	(0.017)
$\Delta(s,t)$	**-0.042 ****	**-0.099 ****	**-0.068 ***	**-0.110 ***
	(0.013)	**(0.032)**	**(0.027)**	**(0.044)**
Discr.-time IV 40-60				
$Y^*_{0,s}$	0,208 **	0,208 **	0,460 **	0,460 **
	(0,003)	(0,003)	(0,004)	(0,004)
$Y^*_{0,t}$	0,268 **	0,340 **	0,061	0,044
	(0,008)	(0,010)	(0,990)	(1,132)
$Y^*_{1,t}(s)$	0,209 **	0,219 **	0,450 **	0,445 **
	(0,003)	(0,003)	(0,054)	(0,071)
$\Delta(s,t)$	**-0,076 ****	**0,153 ****	**0,719**	**0,742**
	(0,010)	**(0,012)**	**(1,745)**	**(1,977)**

** indicates significance at the 1% level, * significance at the 5% level

We use a quadratic specification for the age function in the discrete-time estimations