The Impact of Active Labor Market Programs on the Duration of Unemployment in Switzerland

Rafael Lalive[*], Jan C. van Ours[†] and Josef Zweimüller[‡,§]

July 27, 2006

[*]University of Zurich, CEPR, CESifo, IZA, and IFAU.

[†]Department of Economics, Tilburg University, NL-5000 LE Tilburg, CentER, CEPR and IZA, e-mail: vanours@uvt.nl.

[‡]Corresponding Author. IEW, University of Zurich, IZA, and CEPR. Address: IEW, Bluemlisalpstr. 10, CH-8006 Zurich, Tel. +41-1-634-3724, Fax +41-1-634-4907, e-mail: zweim@iew.unizh.ch.

[§]We thank without implicating Wolfgang Franz, Michael Gerfin, Lorenz Goette, Michael Lechner, Patrick Puhani, Frank Reize, Viktor Steiner, Gerard van den Berg, Rudolf Winter-Ebmer, and two editors as well as three anonymous referees for comments on earlier versions of this paper. We also thank M. Curti and J. Gast at the Swiss State Secretariat for Economic Affairs for help concerning the data and information on institutional details, and Jacek Micuta at the centrale de compensation for access to social security data.

The Impact of Active Labor Market Programs on the Duration of Unemployment in Switzerland

Abstract

This paper evaluates the effects of Swiss active labor market programs on the job chances of unemployed workers. The main innovation is a comparison of two important dynamic evaluation estimators: the "matching" estimator and the "timing-of-events" estimator. Matching and the simple proportional hazard estimator both work if job seekers are selected into programs based on information that is observed but matching is preferable because it does not impose proportionality of the hazard rate. Relying on proportionality, the timing-of-events estimator identifies the causal effect of active labor market programs even when selection on observables fails. We find that both estimators that assume selection on observables agree in finding that temporary subsidized jobs shorten unemployment duration whereas training programs and employment programs do not. In contrast, the timing-of-events estimator suggests that none of the Swiss active labor market programs can shorten unemployment duration.

Keywords: active labor market policy, treatment effect, multivariate duration model

JEL Classification: C14, C41, J64

1 Introduction

The aim of the present paper is to study the impact of active labor market policies (ALMPs) on the duration of unemployment in Switzerland. The new Swiss ALMPs reflect the increasing consensus among policy makers that actively assisting the unemployed in job search is preferable to simply providing them with passive income support. The danger is, so the argument goes, that reliance on passive income support may reduce work incentives and job-search activities and therefore increase the risk of long-term unemployment. ALMPs are seen by many as the key to minimize these risks.

The question how participation in ALMP-measures affects labor market histories of individuals has been the subject of substantial debate. The main problem is that labor market outcomes for participants may be systematically different from non-participants for reasons that are unobservable to the researcher – the selection problem (see Heckman et al., 1999). In Switzerland, like in most European countries, but unlike in the U.S., randomized social experiments are uncommon, so one has to deal with non-experimental data. In theory, several methods can be used to estimate the treatment effects of ALMPs. Each of these methods deals with the selection problem under different assumptions. In the case of unemployment duration as variable of interest two methods are particularly useful. The first one is the method of '*matching*', the second one is the '*timing-of-events*' method.

The main innovation of the present paper is a direct comparison between the timing-of-events approach and the matching approach in estimating the effect of ALMPs on the rate by which unemployed individuals find regular jobs. The method of *matching* is based on the conditional independence assumption. If many variables that influence both labor market outcomes and the selection process are observed, potential outcomes and selection are independent conditional on the observables. The identifying assumption is that, after accounting for many observable variables (including individual's past labor market performance), no unobserved heterogeneity correlated with potential outcomes and program participation is left. Among the many studies that use the matching approach, the studies of Gerfin and Lechner (2002) and Gerfin, Lechner and Steiger (2005) are of interest here as they also evaluate the effect of Swiss ALMPs on unemployment duration. Both studies find that employment programs perform very poorly, vocational training programs show a rather mixed performance depending on the specific subprogram considered, whereas temporary subsidized jobs appears to be successful in terms of increasing the chances on the labor market.[1]

The *timing-of-events* method allows for selection on unobservables in postulating a multivariate mixed proportional hazard model in which both the inflow into an ALMP program and the outflow from unemployment are specified and allowed to interact. The identifying assumption is that these transition processes can be modelled as a multivariate mixed proportional hazard (MMPH) model. The intuition is that, under this assumption, information on the correlations between the unobserved heterogeneity components in the exit from unemployment and the entrance into ALMPs can be obtained from (i) the duration until the program starts and (ii) the duration of unemployment. Because unobserved heterogeneity components are modelled explicitly, the treatment effect is estimated conditional on observed and unobserved variables taking into account that the unobserved variables may influence both processes. The timing-of-events method is a rather new approach and has been applied in only a few previous studies.[2]

In comparing the timing-of-events approach and the matching approach in estimating the effect of ALMPs, we proceed as follows. First, we compare the matching approach with a proportional hazard approach that both rely on conditional independence. We find that the estimated treatment effects are very much the same. While *training programs* and *employment programs* have no effect, temporary *subsidized jobs* have a positive effect on the job finding rate. Second, the timing-of-events approach allows us to introduce potential selectivity in both observable characteristics and unobservable characteristics. If we estimate a MMPH model which allows for selection on unobservable characteristics in addition to observable characteristics none of the treatment effects is positive. So, if unobserved heterogeneity is allowed to influence the inflow into ALMPs the timing-of-events approach and the matching approach find different treatment effects.

The plan of the paper is as follows. In the next section we describe the Swiss labor market policy in more detail. In Section 3 we provide specific information on our data set, a weighted

[1]For a further matching study that also looks at the impact of ALMPs in Switzerland see Prey (2000).

[2]Gritz (1993) considers the impact of training on the employment experience of American youths and Bonnal et al. (1997) study the effect of public employment policies set up in France during the 1980's. Van den Berg et al. (2002) studies the effects of temporary jobs in the Netherlands and both Abbring et al. (2005), Lalive et al. (2005) and Van den Berg et al. (2004) study the effect of benefit sanctions. Two studies closely related to ours are Richardson and Van den Berg (2001) in which the effect of vocational employment training on the transition rate from unemployment to work is investigated and Crépon et al. (2005) who study the effect of counseling programs on unemployment duration and recurrence.

random sample of entrants into unemployment in Switzerland over the four–months period December 1997 to March 1998. Section 4 discusses the modelling of dynamic treatment effects in more detail. The results of our analysis are presented in Section 5. Section 6 concludes.

2 Labor market policy in Switzerland

In 1997 the Swiss government introduced a reform of unemployment insurance that constituted a change away from passive income maintenance towards active measures. The new law obliged the Swiss cantons to supply a minimum number of ALMP-places per year. Economy-wide, these requirements add up to a stock of 25,000 places. This compares to an average stock of unemployment of about 188,000 individuals in 1997 and about 140,000 in 1998.

The new law increased maximum benefit entitlement and, at the same time, created a close link between unemployment benefit entitlement and participation in an active measure. For a newly unemployed the maximum entitlement period is 104 weeks, up from originally 80 weeks.[3] The benefit entitlement period is divided into two different parts. For at most 7 months the job-seeker can receive unemployment benefits, unconditional upon participation in an active measure. For the remaining 17 months unemployment benefits are paid only if the unemployed is willing to participate in a measure.

Employment service staff decides on participation in ALMPs based on subjective evaluation of the job-seekers employment prospects. Individuals are notified about their participation into a program one or two weeks in advance. A job seeker is not allowed to refuse participation once he or she is assigned to participate in an ALMP. Refusal to participate results in withholding of benefit payments for a period of 1 to 30 days.[4]

As mentioned by the OECD (1996), the new Swiss unemployment insurance system is an ambitious one. Compared to other countries the Swiss rules are different in at least two important respects. First, the intervention takes place at a rather early stage of the unemployment spell, after seven months. And secondly, for training courses and employment programs, benefit payments are conditional upon ALMP-participation and this participation *does not* lead to a new benefit entitlement. In contrast, temporary subsidized jobs lead to a new benefit

[3] The above entitlement regulation holds for an individual who has been employed and has contributed to the insurance system for at least 6 within the last 24 months.

[4] See Lalive, van Ours and Zweimüller (2005) for an evaluation of the Swiss sanction system.

entitlement. Note, however, that most individuals enter this program at a rather early stage of the unemployment spell so it is rather unlikely that individuals use this program to acquire prolonged unemployment benefits.[5] The ALMP-measures can be divided into four categories:

1. Basic Training. The job courses usually last 3 weeks and aim at improving the effectiveness of individual job search (how to write application letters, how to behave at job talks) and self-esteem. The computer courses last about 3 weeks and refer to basic word processing and spreadsheet calculation. The language courses last about 2 months and include reading and writing skills. Language courses are more likely to be attended by foreigners but also native Swiss attend these courses frequently.

2. Advanced Training. Vocational training courses last slightly less than two months and provide vocational training in business administration and related areas. Other courses last about 2 months and concern a rather heterogeneous group of course types, including specific computer training, business administration, technical training, courses in the tourism and the health sector.

3. Employment programs. These refer to temporary jobs in the non-profit sector, which last about 5 months. The jobs may be provided by both private sector (NGOs) and public sector (such as communal offices).

4. Subsidized jobs. These are actual low-wage jobs that firms register with the public employment service or that firms offer to an unemployed individual. These jobs are considered to be temporary because the wage in theses jobs is below the official minimum of 67 % of the previous wage (the "reasonability" limit). The individual is still expected to search for a new regular job. It is not possible for firms to reduce the wage payment for such a job in order to benefit from the subsidy.[6]

Table 1

[5] The average elapsed duration at entry is less than 3 months, see Table 1 which is discussed in more detail below.

[6] An unemployed individual who accepts such a low-wage job gets 70 or 80 % of the difference between the previous wage and the wage in the subsidized job as a wage supplement by the unemployment insurance. Note that temporary subsidized jobs are not part of the official ALMP. However, in terms of their set-up and the way they operate there is no reason to exclude them from the analysis. On the contrary, to analyze the effects of policy interventions in full detail it is necessary to include temporary subsidized jobs.

Table 1 presents detailed descriptive statistics on the programs. These statistics, based on the dataset we describe in more detail in the next section, indicate that in terms of participants job training and subsidized jobs are the most important programs. Unemployed workers enter a program after about 3 months of unemployment but the variation is considerable as can be seen from the standard deviation of the elapsed unemployment duration at program entry.

It is worth noting that various programs also differ in terms of hours spent on the program. Training courses typically require weekly hours equivalent to a part-time job, whereas the time-intensity of employment programs are equivalent to a full-time job. Subsidized jobs can be either full-time or part-time. Individuals are required to search for a regular job while attending a program. However, job search requirements are reduced for participants in training courses. One should also note that training courses and employment programs involve costs that go beyond the payment of individual benefits. While training courses and employment programs involve considerable direct costs, this is not the case for subsidized jobs. As the wage subsidy to the temporary jobs amounts to 70 or 80 % of the difference between the previous wage and the new wage, working in a subsidized job increases an individual's income. (The wage plus wage subsidy amounts to more than the unemployment benefit, at most 96 % of the previous wage). As the unemployment benefit amounts to 70 or 80 % of the previous wage, a subsidized job is cheaper in terms of transfer payment from the unemployment insurance system to the individual. As there are no major direct costs, subsidized jobs seem to be a rather cost-effective program.[7]

3 Data

The data set from which we drew our sample, covers all unemployment entrants in Switzerland over the period December 1997 to March 1998 and follows these individuals up to the end of May 1999. These data come from administrative records of the State Secretariat for Economic Affairs (AVAM- and ASAL-data base). Among the 70,445 workers who started an unemployment spell during the above period we concentrate our empirical analysis on a subsample of those workers for whom we could match the information of the AVAM- and ASAL-database with information from social security records (AHV-data). The latter provide detailed information on individuals'

[7]A Swiss study conducted by BASS (1999) estimates that, in the absence of the subsidized job program, total expenditures of the unemployment system as a whole would be 4 percent higher.

earnings and employment histories over the last 10 years prior to their unemployment spells.

We had only limited access to the social security records. These data available to us contains a 50 % random sample of the inflow in December 1997, and a 30 % random sample of the inflow from January 1998 to March 1998. In the analysis in Section 5 we account for this by weighting each observation by the inverse of the probability of being in the random social security sample. The sample on which our empirical estimates are based contains 15,073 job seekers.[8]

Figure 1 shows the transition rates from unemployment to regular jobs and from unemployment to ALMPs.[9] There is a very strong increase in the probability of entering both regular jobs and ALMPs, from about 5 % per month in the first month of unemployment, up to a level of almost 15-20 % per month in the third month of unemployment. Thereafter, the transition rates revert to a level of 5 % per month and remain at that level from month 6 onwards. The patterns of both transition rates are very similar.[10] Thus, the process of finding jobs could be affected by similar factors as the process of finding a suitable active labor market program.

Figure 1 and Figure 2

Figure 2 shows the monthly empirical hazard rates for transitions from unemployment in more detail. The direct transition rate to jobs, labelled as "no ALMP" is identical to the one presented in Figure 1. This baseline exit rate serves to establish the effects of the programs as we discuss in the following section. Defining t_c as the duration until entry into an ALMP we distinguish three groups, those that enter in the first three months "ALMP: $t_c < 3$", those that enter between 3 and 6 months "ALMP: $3 \leq t_c < 6$", and those who enter after 6 months of unemployment "ALMP: $6 \leq t_c$". Compared to the baseline hazard rate, the exit rate of the ALMP-participants is lower initially, but tends to be higher than the baseline hazard rate after a period of at most 4 months. This suggests that capturing the dynamics of the effect of the ALMP on the hazard rate *compared to non-participation* may be important.

[8]We removed all job seekers who were not entitled to unemployment benefits, were re-entering unemployment within a period of two calendar years, were aged younger than 20 years or older than 49 years, were disabled or were foreigners with an asylum seeker or seasonal permit.

[9]The transition rates account for censoring by the Kaplan-Meier method.

[10]Note that with the exception of the employment programs, the temporal pattern of the separate program entry rates is qualitatively similar to the overall ALMP entry rate.

4 Modelling dynamic treatment effects

This section discusses our identification strategy. The logic of our approach is simple. We first discuss two important considerations that arise when a program is dynamically assigned to individuals. Second, we propose two estimators that identify the effects of treatments when treatment assignment is ignorable conditional on the information provided by the data. Whereas the matching estimator just uses the conditional independence assumption, the hazard estimator restricts the hazard rate of the outcome process to follow the proportional hazard restriction. The section finally discusses the possibility that selection is also based on unobservables.

4.1 The start of a treatment

When investigating the effects of ALMPs one has to deal with two questions concerning the start of the program, i.e. the start of the treatment. The first question is whether unemployed individuals anticipated the start of a program; the second question is when the potential effects of the treatment can be expected to occur, right from the start of the program or after the program has ended.

As Abbring and Van den Berg (2003) indicate in the setting of a timing-of-events analysis unemployed individuals are not allowed to anticipate the start of the program a long time in advance. Individuals who anticipate the start of the program may reduce their search intensity prior to the start. In that case the effects of the program are overestimated. We think that anticipation is not a problem in the Swiss case. As discussed in Section 2 job seekers are notified about actual participation only one or two weeks in advance. Therefore, even if they would have wanted to react they did not have a lot of time to react.[11]

The relevant starting date of potential ALMP-effects depends on the nature of the program. In case of training courses, where program duration is established in advance and participants should follow a particular curriculum, it makes sense to start investigating the effects of the program after it has finished. Then, as in Richardson and Van den Berg (2001), the length of time intervals spent in a program is set to zero (i.e. "the calendar time clock is stopped" while participants are in the program). However, in case of subsidized jobs or employment

[11]Furthermore, during our observation period, there was a lack of available ALMP slots so individuals could not anticipate to get into a program eventually (see Lechner and Frölich, 2005). Finally, job seekers were aware of the fact that they could be penalized if they reduced their search efforts in anticipation of program participation.

programs, participants are expected to find a regular job as soon as possible and should accept any suitable job offer. Then, participants can leave the program to take a regular job at any time and stopping the calendar time clock (i.e. disregarding the interval during which the individual is in the program) does not make sense. In other words, whether or not one should investigate the effects on the job finding rate from the start of the program depends on the nature of the program. Note, however, that in both cases there is a locking-in effect. In case of a training program the locking-in effect is exogenous in the sense that participants are expected to finish the program and will not start a new job before the program is finished.[12] In case of a subsidized job or an employment program the locking-in effect is endogenous in the sense that it is determined by the search behavior of the participant. Therefore, for these programs we investigate the treatment effects from the moment they start. To compare the treatment effects across programs we do the same for training courses.

4.2 Selection on observables

4.2.1 Matching estimator

We start our empirical analysis below with a matching estimator that does not restrict the way in which ALMPs affect the exit rate. The estimator recognizes that program participants and non-participants may differ in two important respects. First, program participants may be a selective subset of the population with respect to observables. A meaningful comparison group is therefore balanced with respect to these variables. The second difference arises due to the fact that the timing of program participation is a process with a strong stochastic component.[13] This implies that control individuals must be drawn from the set of individuals that have neither left unemployment nor entered treatment at the moment when the treated individual starts treatment.

In the evaluation, we focus on the effects of the *first* treatment on the duration of unemployment. More precisely, we estimate the effect of a new training "sequence" on the remaining duration of unemployment. If such a sequence consists of participating in two or more ALMPs, information on the occurrence and timing of the second (or third,...) spell is disregarded. The

[12] Note that Richardson and Van den Berg (2001) find a positive treatment effect of vocational employment training only if the time spent in these programs is ignored.

[13] See Fredriksson and Johansson (2003), Gerfin and Lechner (2002), Gerfin, Lechner and Steiger (2005), and Sianesi (2004) for evaluation studies in the random program start setting.

various programs are indexed by $p = 1, ..., 4$.[14] The transition rate from unemployment to program p is assumed to have a proportional hazard specification given by

$$\theta_p(t|x) = \lambda_p(t) \exp(x\beta_p). \tag{1}$$

where θ_p denotes the transition rate from unemployment to a program, t is the elapsed duration of unemployment, and x is a vector of individual and labor market characteristics that determine this transition process.[15] The baseline hazard rate λ_p allows for flexible duration dependence by using a step function

$$\lambda_p(t) = \exp(\Sigma_k \lambda_{pk}\ I_k(t)) \tag{2}$$

where $k = 1, .., 5$ is a subscript for time-intervals and $I_k(t)$ are time-varying dummy variables for the following time intervals: 0-2 months, 3-5 months, 6-8 months, 9-11 months and 12 and more months.[16]

The proposed matching estimator works as follows. In a *first step* we estimate separate transition rates to each of the four programs. This gives an estimate of the transition rate of individual i to each program p: $\widehat{\theta}_p^i(t) = \widehat{\lambda}_p(t) \exp(x_i \widehat{\beta}_p)$.

In the *second step* we select, for each participant i in program p, the "nearest neighbor" in terms of the transition rate to program p. Note that the transition rate to the program is the instantaneous propensity score. The set of potential nearest neighbors consists of observations that are still "at risk" of entering program p as their first program. In other words, these are observations that have neither entered an ALMP nor accepted a regular job before observation i entered program p. Let T_u denote the random variable that characterizes the duration of unemployment until the individual finds a regular job, and let T_p denote the random variable that characterizes the duration of unemployment until the start of program p. We denote by $\tilde{T}$ the random variable that characterizes the duration until the individual either finds a regular job or starts her first treatment, so $\tilde{T} = \min(T_u, T_1, ..., T_4)$. And we denote by $\tilde{t}$ a particular realization

[14] $p = 1$ indexes basic training, $p = 2$ advanced training, $p = 3$ employment programs and $p = 4$ subsidized jobs.

[15] Specifically, these observables are gender, marital status, number of dependents, age, residence permit (applies to non-Swiss residents), mother tongue, skill level, position in the previous firm, type of job required (in same industry, part-time), previous industry, previous occupation, previous wage, duration of previous job, recent labor market history (1995-1997), distant labor market history (1988-1994), inflow period, the unemployment rate in the canton of residence of the job seeker in the month prior to entering unemployment (time-invariant), proportion of unemployed in ALMP, voting in a 1997 referendum on benefit cuts, and employability.

[16] In a sensitivity analysis, we use a very flexible baseline hazard which is allowed to shift every month.

of $\tilde{T}$. The set of potential control observations for individual i with $T_p^i = \tilde{t}^i$ (somebody whose first program is p) is given by $A_i \equiv \{j|\tilde{t}^j > \tilde{t}^i, j \neq i\}$. The "nearest neighbor" is the observation $j \in A_i$ that minimizes $abs(\widehat{\theta}_p^j(\tilde{t}^i) - \widehat{\theta}_p^i(\tilde{t}^i))$.[17]

The *final step* involves estimating the counterfactual survivor function of the treated observations by using the information provided by control observations. There are two important considerations in this step. First control observations may be treated in the future.[18] Thus, for control observations the remaining duration of unemployment is given by $\min(T_u^j - T_p^i, T_1^j - T_p^i, ..., T_4^j - T_p^i)$, i.e. the time remaining in unemployment without participating in any of the four programs. The counterfactual survivor function can be recovered from information on this remaining duration of control observations, treating as right-censored all observations going to an ALMP. Second, the asymptotic bias may dominate the asymptotic variance in nearest-neighbor matching (Abadie and Imbens, 2002). This problem can be addressed by estimating the counterfactual transition rate from unemployment to regular jobs conditional on the observed characteristics of the matched controls. This estimate of the counterfactual transition rate can then be used to simulate the counterfactual survivor function for the treated population using the observed characteristics of the treated.

The central identifying assumption that justifies the matching estimator is that conditional on observed characteristics of the individuals assignment to program p is independent of the potential remaining duration without the program. This conditional independence assumption has been justified by Gerfin and Lechner (2002) by the fact that the Swiss unemployment insurance register is extremely rich in terms of observed characteristics. In particular, the data contain information on employability, a subjective caseworker assessment of the likely labor market prospects of the job seeker filled out at the start of the unemployment spell.[19] Fredriksson and Johansson (2003) show that when the conditional independence assumption holds, the effect of treatment on the treated can be identified by contrasting the outcomes of individuals who are treated at $\tilde{t}$ to individuals who have not been treated until $\tilde{t}$ who have the

[17]Similar to Sianesi (2004) we impose a caliper of .01 in determining the nearest neighbor.

[18]This has led to focusing on the effect of "being treated now vs. being treated in the future" (Fredriksson and Johansson, 2003; Sianesi, 2004).

[19]In addition, we use information on gender, marital status, number of dependents, age, residence permit (applies to non-Swiss residents), mother tongue, skill level, position in the previous firm, type of job required (in same industry, part-time), previous industry, previous occupation, previous wage, duration of previous job, recent labor market history (1995-1997), distant labor market history (1988-1994), inflow period, unemployment rate, proportion of unemployed in ALMP, voting in a 1997 referendum on benefit cuts to control for differences in labor market prospects between ALMP participants and non-participants.

same propensity to enter treatment at time $\tilde{t}$. Note that the propensity to enter treatment at time $\tilde{t}$ is identical to the ALMP entry hazard rate $\theta_p(\tilde{t}|x)$.

4.2.2 Proportional hazard model

Alternatively, we use a proportional hazard model to identify the treatment effects of the various programs. Let

$$D_p(t) = I[t > \tilde{t} \cup \tilde{t} = T_p] \tag{3}$$

be the indicator function that, after elapsed duration t, the individual has already entered his or her first program and that this is program p. This defines 4 treatment indicators $D_1, ..., D_4$, one for each program.

Note that because we analyze the treatment effects of the four programs separately, it is essential to censor unemployment spells for job seekers leaving unemployment for a program that is not being studied. For instance, when studying basic training the unemployment spell is recorded as right censored for all job seekers who enter advanced training, employment programs, or subsidized jobs at the time when they enter those programs. The proportional hazard estimator postulates that the transition rate from unemployment to regular jobs is

$$\theta_u(t|x, D_p(t)) = \lambda_u(t) \exp(x\beta_u + \delta_p(t, \tilde{t}) D_p(t)). \tag{4}$$

where $\theta_u(t|\cdot)$ is the transition rate from unemployment to a regular job at elapsed duration of unemployment t conditional on individual characteristics x and the treatment indicator $D_p(t)$. The treatment effects are specified as $\delta_p(t, \tilde{t}) = \Sigma_k \delta_{pk} I_k(t - \tilde{t})$ where the δ_{pk} are parameters to be estimated, and $I_k(t - \tilde{t})$ are dummy variables that vary with time since start of treatment $t - \tilde{t}$ for the intervals 0-2 months, 3-5 months, 6-8 months, and 9 months and longer. The baseline hazard $\lambda_u(t)$ is again allowed to vary with the elapsed duration of unemployment in the same way as the program entry hazard rate. In separate estimates, the model estimates 4 treatment effect vectors δ_p, one for each program. We keep the specification flexible and allow treatment effects to vary over time.

The proportional hazard estimator identifies the effect of ALMPs on the duration of unemployment under two conditions. First, it assumes that conditional on observables x, participation in the ALMP is not informative on unemployment duration without the program, i.e. that selection is based on observables. This is the assumption that is also required for the matching estimator. In addition, the proportional hazard estimator imposes a particular functional form on the hazard rate. The assumption is that the characteristics x of the individuals shift the hazard rate in a proportional manner irrespective of the time elapsed since the start of the spell.

4.2.3 Comparing the two methods

The focus of our evaluation is the causal effect of treatment p on the remaining duration of unemployment after the start of the program. Remaining duration of unemployment is given by $T_r^p \equiv T_u^p - \widetilde{T}$, where T_u^p is total duration of unemployment with program p, and $\widetilde{T}$ is the duration of unemployment until the first program starts. The counterfactual is $T_r \equiv T_u - \widetilde{T}$, that is, the remaining duration of unemployment without the program.

We compare the results of the matching estimator and the proportional hazard estimator with respect to the effect of treatment p on those treated with program p, i.e. $E(T_r^p - T_r|\widetilde{T} = T_p)$. The effect of treatment on the treated is useful in assessing whether program p has achieved the goal to foster re-entry of job seekers into regular jobs. Note, however, that the effect of program p should not be compared to the effect of program p' because the program p applies to a different subpopulation than the effect of program p'. An alternative parameter is the average effect of treatment, that is, the effect of treatment on the average job seeker. This parameter is useful in discussing the issue of whether the program should be extended to the entire population of job seekers. We focus on the effect of treatment on the treated because this parameter is crucial in the ex post evaluation of the question whether active labor market programs are helpful in placing job seekers that were affected by the programs into jobs.

Since both T_r^p and T_r are positive random variables, the effect of treatment on the treated can be represented as follows

$$E(T_r^p - T_r|\widetilde{T} = T_p) = \int_0^\infty (S_r^p(t|\widetilde{T} = T_p) - S_r(t|\widetilde{T} = T_p))dt \qquad (5)$$

where $S_r^p(t|\widetilde{T} = T_p)$ is the survivor function of remaining duration with treatment p, i.e. $S_r^p(t|\widetilde{T} = T_p) = 1 - \Pr(T_r^p > t|\widetilde{T} = T_p)$, and $S_r(t|\widetilde{T} = T_p)$ is the survivor function of the counterfactual remaining duration without treatment p. Note that right-censoring of the remaining duration of unemployment implies that the effect of treatment on the treated can not be recovered from the data. Instead, we base our comparison of the results on the difference in the survivor function with treatment and the counterfactual survivor function without treatment in the first 12 months after the start of the treatment.[20]

The *matching estimator* provides a matched set of treated and control observations. We estimate the unconditional (with respect to x) survivor function with treatment and the unconditional counterfactual survivor function in three steps. First, we estimate the conditional hazard of leaving unemployment for regular jobs using maximum likelihood. Note that we estimate the conditional counterfactual hazard of leaving unemployment without treatment using the matched control sample. In the second step, we use the resulting estimates to simulate both the conditional survivor function with treatment $S_r^p(t|x, \widetilde{T} = T_p)$, and the conditional counterfactual survivor function without treatment $S_r(t|x, \widetilde{T} = T_p)$ for each person in treatment p. In the third step, we average the conditional survivor functions to estimate the unconditional survivor function with treatment $S_r^p(t|\widetilde{T} = T_p) \equiv E[S_r^p(t|x, \widetilde{T} = T_p)|D_p = 1]$, and the unconditional counterfactual survivor function without treatment $S_r(t|\widetilde{T} = T_p) \equiv E[S_r(t|x, \widetilde{T} = T_p)|D_p = 1]$. This three step procedure addresses the problem that nearest neighbor propensity score matching methods may be inconsistent (Abadie and Imbens, 2002).

The *proportional hazard estimator* gives the conditional (on observables x and program entry times $\widetilde{t}$) remaining duration survivor function with treatment $\widehat{S}_r^p(t|x, \widetilde{t}) = \exp(-\int_0^t \widehat{\theta}_u(\widetilde{t} + z|x, D_p(\widetilde{t} + z))dz)$.[21] The conditional survivor function without treatment is obtained by imposing non-participation, i.e. $D_p(\widetilde{t} + z) = 0$ everywhere. The unconditional survivor curves are obtained by taking the average with respect of the distribution of x and program entry times $\tilde{t}$ in the treated population of the corresponding conditional survivor functions.[22]

The comparison of the two estimators is based on the difference in the unconditional survivor

[20] Note that the integral with respect to time since start of this difference gives the "effect of treatment on the treated in the first 12 months after start." We restrict attention to the first 12 months after start due to right censoring. Since the average time until the first program starts is between 3 and 4 months and since the observation window covers at least 14 months (for those entering end of March 1998), censoring is unlikely to affect the results regarding the first 12 months after the start of the program.

[21] Note that $D_p(\widetilde{t} + z) = 1$ since $\widetilde{t}$ is the date of program entry.

[22] A second alternative to using bootstrap standard errors is use the asymptotic distribution of the proportional hazard model parameters to calculate the asymptotic standard errors of the treatment effect in the proportional hazard model. This strategy tends to give smaller standard errors because the parametric model is more efficient. Nevertheless, we find that our main conclusions regarding the comparison between the matching estimator and the proportional hazard estimator are robust to using standard errors due to the proportional hazard model.

curves.[23] This difference should be negligible if (i) the unemployment exit rate indeed has a proportional structure, and (ii) the proportional hazard model is sufficiently flexible to capture treatment effect heterogeneity and the dynamics of the treatment effect. Note that even if the empirical results suggest that the difference in the matching and proportional hazard estimates is not statistically different from zero it does not necessarily follow that the proportionality restriction is valid. It appears possible to construct examples where proportionality is violated but the semi-parametric matching method and the proportional hazard estimator nevertheless provide similar estimates. We nevertheless believe that it is instructive to perform this analysis because it documents how strongly the important proportionality assumption is affecting results.

Inference is based on the variability of the difference between the effect of treatment on the treated survivor curve according to matching and the effect of treatment on the treated survivor curve according to the proportional hazard model in 250 sub-samples of the original dataset. While the asymptotic distribution of the proportional hazard estimator are well understood, we are not aware of asymptotic results for propensity score matching estimators that account for variability of the first stage (see the survey by Imbens 2004). Note that bootstrapping leads to biased inference on the asymptotic variance of the matching estimator (Abadie and Imbens, 2004). We therefore use subsampling. Politis and Romano (1994) show that subsampling works if the sampling distribution of the difference in survivor curves converges weakly to the underlying population distribution and the ratio between the sub-sample size b and the sample size n converges to zero as n tends to ∞. Theoretical considerations regarding the choice of b are difficult and beyond the scope of this paper. In this application, we choose b so large that the probability is high that all models can be calculated in subsamples.[24]

4.3 Allowing for selection on unobservables

The third estimator relies strongly on the assumption of proportionality but it relaxes the assumption of conditional independence. Specifically, this estimator is based on the following mixed proportional hazard specification for the transition rate unemployment to a regular job

$$\theta_u(t|x, D_p(t), v_u) = \lambda_u(t) \exp(x\beta_u + \delta_p(t, \tilde{t}) D_p(t)) v_u. \tag{6}$$

[23]Note that in performing this comparison, we restrict attention to the set of participants in program p, for whom we can find a "nearest neighbor" according to the matching protocol.

[24]Specifically, we fix $b = int(n^{99/100})$ which is $b = 13,690$.

The term v_u captures heterogeneity that is unobserved to the researcher and is allowed to be correlated with corresponding heterogeneity terms v_p in the transition rate from unemployment to program p, and and v_c in the process that characterizes endogenous right censoring when job seekers exit unemployment for other active labor market programs. The model for the transition rate from unemployment to program p is

$$\theta_p(t|x, v_p) = \lambda_p(t) \exp(x\beta_p) v_p, \tag{7}$$

and the model for entry into other programs – the endogenous right censoring process – is

$$\theta_c(t|x, v_c) = \lambda_c(t) \exp(x\beta_c) v_c, \tag{8}$$

and the unknown joint distribution of the heterogeneity terms is denoted by $G(v_u, v_p, v_c)$.

Abbring and Van den Berg (2003a) prove that the model consisting of (6) and (7) is identified. Because entry into other active labor market programs is likely to be endogenous, we add the third censoring process (8) to the basic 'timing-of-events' model. The treatment effect in this extended MMPH model is identified. The identification proof in Abbring and Van den Berg (2003a, p. 550) has two parts. The first part notes that a model that censors the outcome process at the time of entry into program p is a basic and well-understood competing risks model with unobserved heterogeneity. This model is identified regardless of the number of processes (Abbring and van den Berg, 2003b). The second part of the proof shows that the treatment effect is identified. This result does not depend on the number of competing risks process in the MMPH model. It follows that the model consisting of the processes (6), (7), and (8) is identified.[25]

Estimating the model requires specification of the joint distribution of the heterogeneity terms $G(v_u, v_p, v_c)$. We follow the standard approach in the literature of approximating the unknown joint distribution by means of a discrete distribution using non-parametric maximum likelihood (NPMLE). We assume G to be a multivariate discrete distribution of unobserved heterogeneity. Work by Heckman and Singer (1984) suggests that discrete distributions can approximate any arbitrary distribution function G. We assume that each transition rate has

[25]We thank Gerard van den Berg for pointing this out to us.

two points of support – $(v_{u,a}, v_{u,b})$, $(v_{p,a}, v_{p,b})$, $(v_{c,a}, v_{c,b})$ – so the joint distribution therefore has eight points of support.

The MMPH model relaxes the assumption of conditional independence of the potential durations from program participation status. Note, however, that this generality comes at a cost. First, it is necessary to specify a functional form in which heterogeneity enters the hazard rate.[26] Second, in single spell data, we have to assume that unobserved heterogeneity is independent of the observables x. Third, as with the PH estimator, the assumption of proportionality needs to hold. If these restrictions hold, a comparison between the PH and the MMPH estimator allows investigating the extent to which the assumption of "selection on observables" affects the estimated effect of ALMPs on unemployment duration.

5 Estimation results

5.1 Accounting for selection on observables

We present results of the matching estimator in Figure 3. The vertical axis measures the differences between the survivor function of the treated and the counterfactual survivor function estimated from matched control observations. For basic training and for employment programs, this difference is positive over almost the complete year after the program start. Even one year after program has started the difference is close to zero or even slightly positive. Taken together this means that basic training and employment programs prolong the duration of unemployment. Both advanced training and subsidized jobs also tend to prolong unemployment initially during the first 4 months (subsidized job) to 6 months (advanced training) probably due to a locking-in effect. As time passes, however, there is a clear negative difference between the survivor function with treatment compared to the counterfactual. This difference is statistically significant 6 months after a subsidized job has started. The difference remains insignificant for advanced training throughout the first year after the program started. This suggests that in the medium to long run advanced training and subsidized jobs can lead to a reduction in average unemployment duration.

Figure 3

[26]Note that the above specification is more restrictive, however, than some of the models discussed in Abbring and van den Berg (2003). For instance, the treatment effect is allowed to vary with respect to observables and unobservables in Abbring and Van den Berg (2003).

The results presented in Figure 3 can be compared to the results in Gerfin and Lechner (2002) and Gerfin, Lechner and Steiger (2003). In these papers, the difference in the survivor curves are also increasing at early durations reaching a maximum after 3 to 5 months after the program started and then start to decline. In quantitative amount the effects are somewhat different, though. This may be due to two reasons. First, our sample differs from the one used in Gerfin and Lechner (2002). The latter use a stock sample, whereas our sample is an inflow sample. Second, our control group consists of individuals that are *not yet* treated but may be treated at a later stage of the unemployment spell (in which case the information on the remaining duration after program start is taken as censored). In Gerfin and Lechner (2002) the control group consists only of individuals that are *never* treated. Note, however, that in qualitative terms the dynamic patterns of the treatment effect is very similar. We observe an increase in the difference in survivor rates between treatment and control group at early remaining durations, and the opposite pattern at later durations. Moreover, also in Gerfin and Lechner (2002) subsidized jobs seem to be quite successful.

The second estimator that can be used to identify the causal effects when selection into the programs is based on observables is the proportional hazard estimator. Table 2 shows how the four programs affect the transition rate from unemployment to regular jobs as a function of time elapsed since the program started. There is a significant reduction in the transition rate from unemployment to regular jobs in the first 3 months (0 to 2 months) after the program started for all programs except for the subsidized jobs. This "locking-in-effect" is strongest for employment programs implying a reduction of the hazard rate by 53 % $(= 100(exp(-.765) - 1)$. The training programs are characterized by somewhat weaker "locking-in-effects" on the order of 19 % for basic training and 24 % for advanced training. Exits from unemployment to regular jobs are, however, already slightly higher for the treated compared to the counterfactual situation 3 to 5 months after the program starts for all programs except for basic training programs. The improvement in the hazard rate is, however, only significant for subsidized jobs. During 6 to 8 after their start, all Swiss active labor market programs are shown to improve the job chances of participating job seekers. Only for basic training, the positive effect is not significantly different from zero. When 9 and more months have elapsed all programs significantly improve job chances of job seekers.

It is interesting to know whether the initial negative effect of most of the programs is more than compensated later on, i.e. whether the net program effect is positive. One way to investigate the net program effect in the context of a proportional hazard model is to use time-invariant treatment effects (Panel B in Table 2). Proportional hazard models with a time-invariant treatment effect indicate that the net effect is significantly negative for basic training and employment programs. The net effect point estimate is positive but not significantly different from zero for advanced training. Subsidized jobs are the only program with a statistically significantly positive effect on exits from unemployment to regular jobs. The results concerning the subsidized jobs imply that on average these jobs increase the regular job finding rate with 9.4%.

Table 2

Figure 4 compares the results due to the matching estimator with the result due to the proportional hazard estimator. This comparison is important. The matching estimator is just identified if selection into treatment is conditionally independent of potential outcomes. The proportional hazard estimator also requires exogenous participation but, in addition, also imposes a proportional structure on the unemployment exit rate. A comparison of matching results and proportional hazard results thus assesses the *robustness* of our findings to the imposing proportionality of the hazard rate. Figure 4 reports the difference in the causal effect according to the proportional hazard estimator and the causal effect according to the matching estimator. A positive number thus indicates that the proportional hazard estimator is more pessimistic regarding the effects of Swiss active labor market programs on unemployment duration. Figure 4 also reports the 95 % confidence interval on the difference in causal effects estimated by subsampling (see section 4). Figure 4 shows that the results for the proportional hazard estimator are basically *identical* to the results for the matching estimator in a statistical sense. There is no statistically significant divergence of results for any of the four programs considered. Figure 4 thus shows that if conditional independence is valid, the results are not sensitive to imposing a proportional structure on the hazard rate.[27]

Figure 4

[27]Note that this does not imply that the proportional structure is correct. It merely implies that the proportional structure does not bias results in a statistically significant way. Moreover, Figure 4 also does not allow investigating whether assuming proportionality for observed *and unobserved* characteristics biases results.

5.2 Allowing for unobserved heterogeneity

We study the treatment effects of the programs in more detail by introducing unobserved heterogeneity into the analysis and estimate MMPH models. Table 3 reports the estimated treatment effects

Table 3

As shown, in each of the estimated models unobserved heterogeneity is identified although the number of masspoints depends on the program investigated. For instance, there are four masspoints for basic training. Conditional on observed characteristics and elapsed duration there is a group of unemployed individuals consisting of 93.0% of the sample that have a high exit rate to a regular job, a high exit rate to a course and a high exit rate to other programs – the censoring rate. The other groups of 3.3%, 1.8% and 1.9 % have different combinations of transitions rates but the shear size of the first group implies that there is a positive correlation between the unobserved components of the job finding rate and the transition rate to courses. There could be several reasons for such a positive correlation. It could result from the incentives of caseworkers. In order to have a favorable placement record, caseworkers may send those unemployed with the highest chances of getting a regular job into basic training. It could also be the case that individuals with the better chances to get a regular job are better motivated to do a course for some intermediate period.

The number of mass points identified ranges from three in the model with advanced training to six in the model with employment programs while in the case of the model with subsidized jobs three mass points are identified. For each of the models there is a predominant positive correlation between the exit rate to regular jobs and the exit rate to the particular program. If these positive correlations are not accounted for, the treatment effects will be overestimated. Indeed, as shown in Table 3, once we allow for unobserved heterogeneity the treatment effects of all programs are either negative or not statistically different from zero.

Panel B in Table 3 reports the net effect of these programs on exits from unemployment to regular jobs. This net effect is significantly negative for basic training, advanced training, and employment programs. The net effect is not statistically different from zero for subsidized jobs.

To investigate the robustness of our results we perform a variety of sensitivity analyses, one of which is shown in Table 4.[28] Recall that the main result for the jobs was obtained in a trivariate MPH model that allows for a shift in the baseline hazard rate after 3, 6, 9, and 12 months respectively. Table 4 shows that the relevant parameter estimates of trivariate MPH models that allow for a shift in the baseline hazard rate after every month, i.e. after 1, 2, ..., 17 months. Results for the model with a flexible specification of the baseline hazard are similar to the baseline results. Changing the specification of the baseline hazard does not affect the estimates of the underlying heterogeneity distribution. Moreover, the flexible baseline hazard model also indicates that the treatment effects are negative or not statistically significant from zero.

Table 4

Table 5 allows for a time-of-entry effect in the causal effect of training programs.[29] From a statistical point of view it may be that unobserved heterogeneity is capturing functional form misspecification in the baseline model. Suppose that program effects vary with time of entry in the sense that the causal effect of a program is worse when individuals enter the program late in the unemployment spell. Neglecting such a time-of-entry effect then might lead to wrongly identifying unobserved heterogeneity because there is a group of job seekers (entering late) with low exit rates and low program entry rates and another group of job seekers (entering early) with a high exit rate and a high program entry rate. It is therefore important to assess the sensitivity of our results to allowing for time-of-entry effects. Table 5 shows that time-of-entry effects matter for all programs except for the advanced training courses. The results indicate that programs work better when job seekers enter early rather than late. For instance, entering a program one month later is shown to decrease the effect of basic training program 5.5 percentage points ($= 100(exp(-.057) - 1)$). Nevertheless, the main conclusion from the baseline model remains unaffected. All Swiss active labor market programs either decrease the exits from unemployment

[28] In addition to this we estimated MMPH models for sub-programs. We did separate estimates for job courses, language courses, computer courses, further vocational training, other courses, public employment programs, and private employment programs. This did not change our main conclusions. As an alternative to the discrete distribution of unobserved heterogeneity we tried using a multivariate log normal distribution of unobserved heterogeneity. However, we were unable to find any improvement in the log likelihood compared to the model that does not allow for unobserved heterogeneity. Apparently the multivariate log normal specification is too restrictive.

[29] We are grateful to a referee for raising this issue.

to regular jobs or their effects are not significantly different from zero because there appears to be *genuine* unobserved heterogeneity in exits to regular jobs, entry into the program that is being studied, and entry into other programs (endogenous right-censoring).

Table 5

6 Conclusions

This paper discusses the effect of ALMPs on the duration of unemployment in a dynamic evaluation context. In the empirical analysis we discuss in detail to what extent the functional form assumption of the proportional hazard model and the assumption of conditional independence may affect the evaluation results.

The empirical results of our paper come in three parts. First, we use a matching method presenting the treatment effect results in the form of graphs. Though the set-up of the matching estimator is different from the one in previous studies on the effectiveness of Swiss labor market policies the results are very much the same. Second, we use a proportional hazard model with time-varying treatment effects. Both approaches lead to the same conclusion that the program of subsidized jobs is the most promising program in terms of their positive effects on the transition rate from unemployment to regular jobs. Third, we estimate a bivariate MPH-model where regular jobs and ALMPs are competing destinations. In the context of this model the treatment effect can be estimated accounting for selectivity both due to observed and due to unobserved characteristics. We conclude that after allowing for selectivity even the treatment effect of subsidized jobs fades away. The reason is that the unobserved characteristics in the job finding rate and the program entrance rate are positively correlated.

From a research point of view our main result is that the matching approach and the timing-of-events approach generates different treatment effects once we allow unobserved heterogeneity to influence the inflow into ALMPs. It is difficult to compare both methods directly as neither of them has a clear economic interpretation and the identifying assumptions are not nested. The method of matching is based on the conditional independence assumption, i.e. the assumption that potential outcomes and selection into programs are independent conditional on the observables. If this assumption is valid, the method of matching is to be preferred to other methods since it is just-identified. In the timing-of-events approach it is possible to relax this assumption and allow unobserved heterogeneity to affect the selection process. However this comes at a cost since it requires assumptions with respect to functional form and independence between unobservables and observables.

From a policy point of view our main result is that the introduction of unobserved heterogeneity substantially affects the estimated treatment effect. This implies that further and more detailed information regarding how job seekers are selected into programs is crucial before policy recommendations can be made.

References

[1]

[2] Abadie, Alberto and Guido W. Imbens (2002), Simple and Bias-Corrected Matching Estimators for Average Treatment Effects, NBER technical working paper No. 283.

[3] Abadie, Alberto and Guido W. Imbens (2004), On the Failure of the Bootstrap for Matching Estimators, Mimeo, Harvard University.

[4] Abbring, J.H., Van den Berg, G.J., and Van Ours, J.C. (2005), The Effect of Unemployment Insurance Sanctions on the Transition Rate from Unemployment to Employment, *Economic Journal*, 115, 602-630.

[5] Abbring, J.H, and Van den Berg, G.J. (2003a), The Non-Parametric Identification of Treatment Effects in Duration Models, *Econometrica*, 71, 1491-1518.

[6] Abbring, J.H, and Van den Berg, G.J. (2003b), The Identifiability of the Mixed Proportional Hazard Competing Risks Model, *Journal of the Royal Statistical Society*, Series B, 65, 701-710.

[7] BASS (1999), Evaluation der Regelung des Zwischenverdienstes in der Schweiz, seco Publikation Arbeitsmarktpolitik Nr.1.

[8] Bonnal, L., Fougère, D., and Sérandon, A. (1997), Evaluating the Impact of French Employment Policies on Individual Labor Market Histories, *Review of Economic Studies*, 64, 683-713.

[9] Crépon, B., Dejemeppe, M. and M. Gurgand (2005), Counseling the Unemployed: Does It Lower Unemployment Duration and Recurrence?, IZA Discussion Paper No. 1796, IZA, Bonn.

[10] Fredriksson, P., and Johansson, P. (2003), Program Evaluation and Random Program Starts, *mimeo*, IFAU.

[11] Gerfin, M., and Lechner, M. (2002), Microeconometric Evaluation of Active Labor Market Policies in Switzerland, *Economic Journal*, 112, 854-893.

[12] Gerfin, M., Lechner, M., and Steiger, H. (2005), Does subsidized temporary employment get the unemployed back to work? *Labour Economics*, 12, 807-835.

[13] Gritz, M. (1993), The Impact of Training on the Frequency and Duration of Employment, *Journal of Econometrics*, 57, 21-51.

[14] Heckman, J.J., and Singer, B. (1984), A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, 52, 271-320.

[15] Heckman, J.J., LaLonde, R.J., and Smith, J.A. (1999), The Economics and Econometrics of Active Labor Market Programs, in: Ashenfelter, O., and Card, D. (eds.) *Handbook of Labor Economics*, volume III, North-Holland.

[16] Imbens, G. W. (2004), Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review, *Review of Economics and Statistics*, 86(4), 4-29.

[17] Kalbfleisch, J. D. and R. L. Prentice (2002), *The Statistical Analysis of Failure Time Data*, 2nd ed. New York: John Wiley & Sons.

[18] Lalive, R., Van Ours, J.C., and Zweimüller, J. (2005), The Effect of Benefit Sanctions on the Duration of Unemployment, *Journal of the European Economic Association*, 3, 1-32.

[19] Lechner, M. (2002), Some practical issues in the evaluation of heterogeneous labor market programmes by matching methods, *Journal of the Royal Statistical Society - Series A*, 165/1, 59-82.

[20] Lechner, M. and Frölich, M. (2005), Regional Treatment Intensity as an Instrument for the Evaluation of Labour Market Policies, IZA Discussion Paper No. 1095.

[21] OECD (1996), *Labor Market Policies in Switzerland*, Paris.

[22] Politis, D. N. and Romano, J.P. (1994), A general theory of large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics* 22, 2031-2050.

[23] Prey, H. (2000), Evaluation of Training Programs in St. Gallen, Switzerland. *Swiss Journal of Economics and Statistics*, 136, 417-432.

[24] Publicom AG, M.I.S. Trend SA (1998), Qualitätsbeurteilung der Regionalen Arbeitsvermittlungszentren (RAV) und der Arbeitsmarktlichen Massnahmen, Schlussbericht, Studie im Auftrag des Bundesamtes für Wirtschaft und Arbeit (BWA), Bern.

[25] Richardson, K., and Van den Berg, G.J. (2001), The Effect of Vocational Employment Training on the Individual Transition Rate from Unemployment to Work, *Swedish Economic Policy Review*, 8, 2, 175-213.

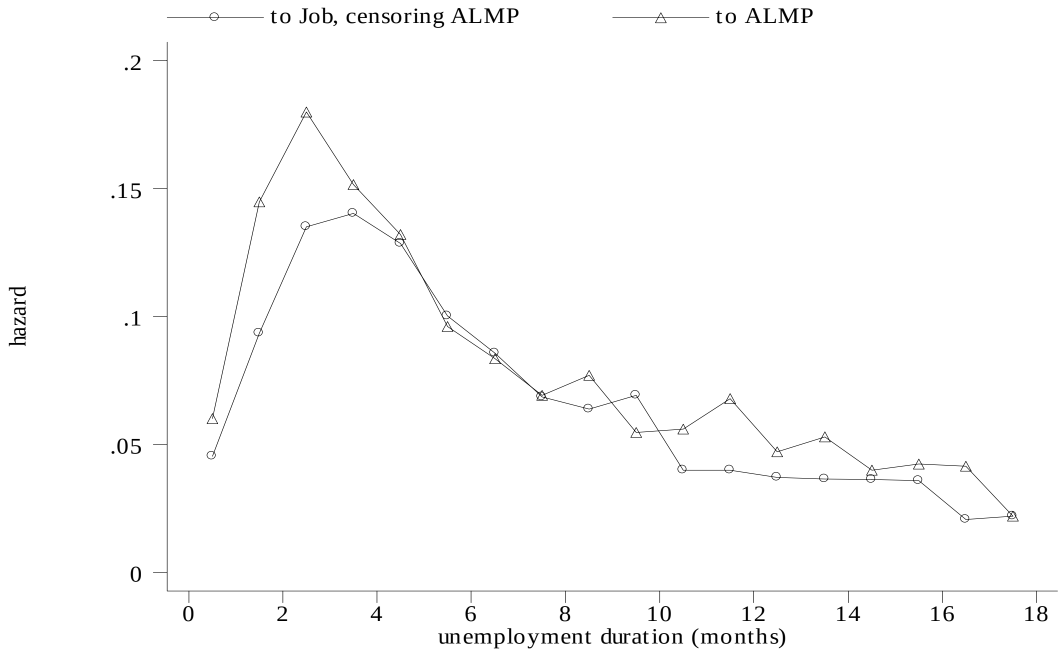
[26] Sianesi, B. (2004), An Evaluation of the Swedish System of Active Labour Market Programmes in the 1990s, *Review of Economics and Statistics*, 86(1), 133 - 155.

[27] Van den Berg, G.J. (2001), Duration Models: Specification, Identification, and Multiple Durations, in: Heckman, J.J., and Leamer, E. (eds.) *Handbook of Econometrics*, Volume V, North-Holland.

[28] Van den Berg, G.J., Holm, A., and Van Ours, J.C. (2002), Do Stepping-stone Jobs exist? Early Career Paths in the Medical Profession, *Journal of Population Economics*, 15, 647-665.

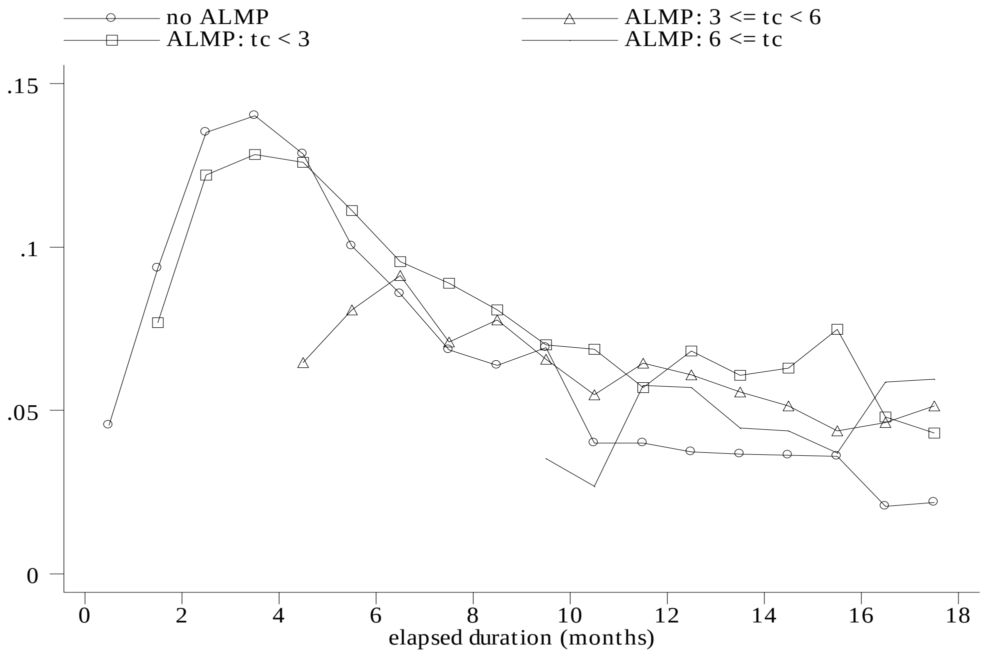
[29] Van den Berg, G.J., Van der Klaauw, B., and Van Ours, J.C. (2004), Punitive Sanctions and the Transition Rate from Welfare to Work, *Journal of Labor Economics*, 22, 211-241.

Figure 1:
Transition rate to Job and to ALMP



Source: Own calculations, based on Swiss unemployment register data.

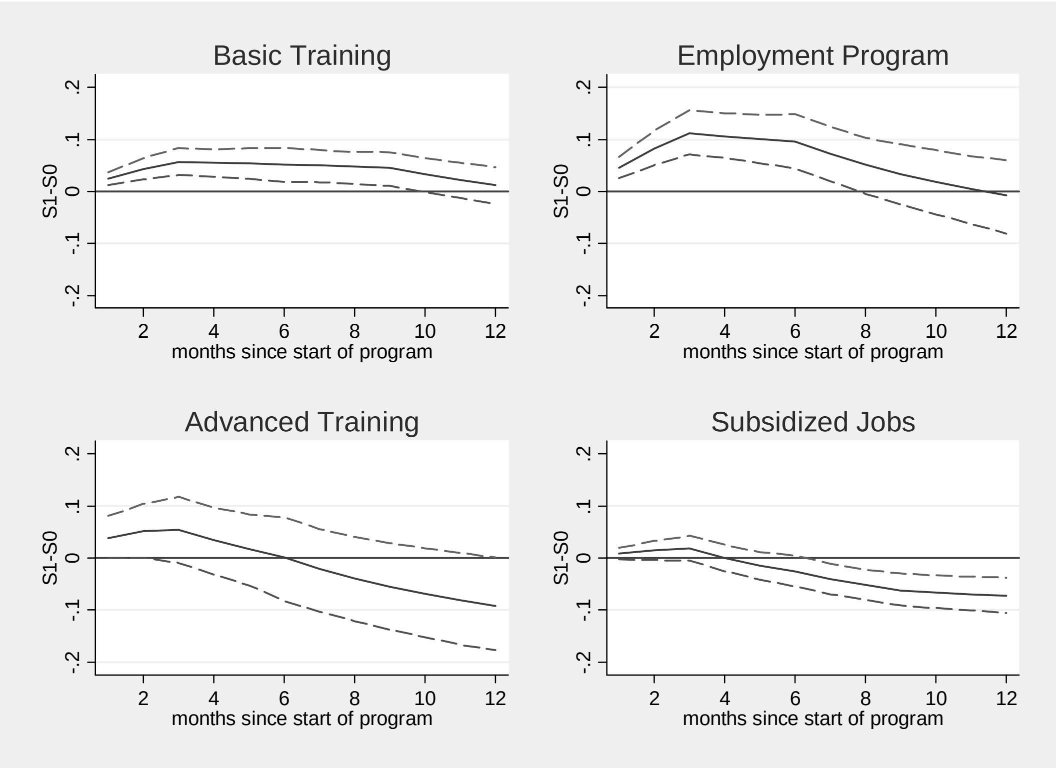
Figure 2:
Transition rate to regular jobs, by treatment status



Notes: tc refers to duration until entry. The average duration until entry is 1.7, 4.1, and 9.3 months for the three groups, respectively. no ALMP refers to the transition rate to jobs treating exits to ALMP as censored.

Source: Own calculations, based on Swiss unemployment register data.

Figure 3:
The Effects of Active Labor Market Programs
Matching Estimator

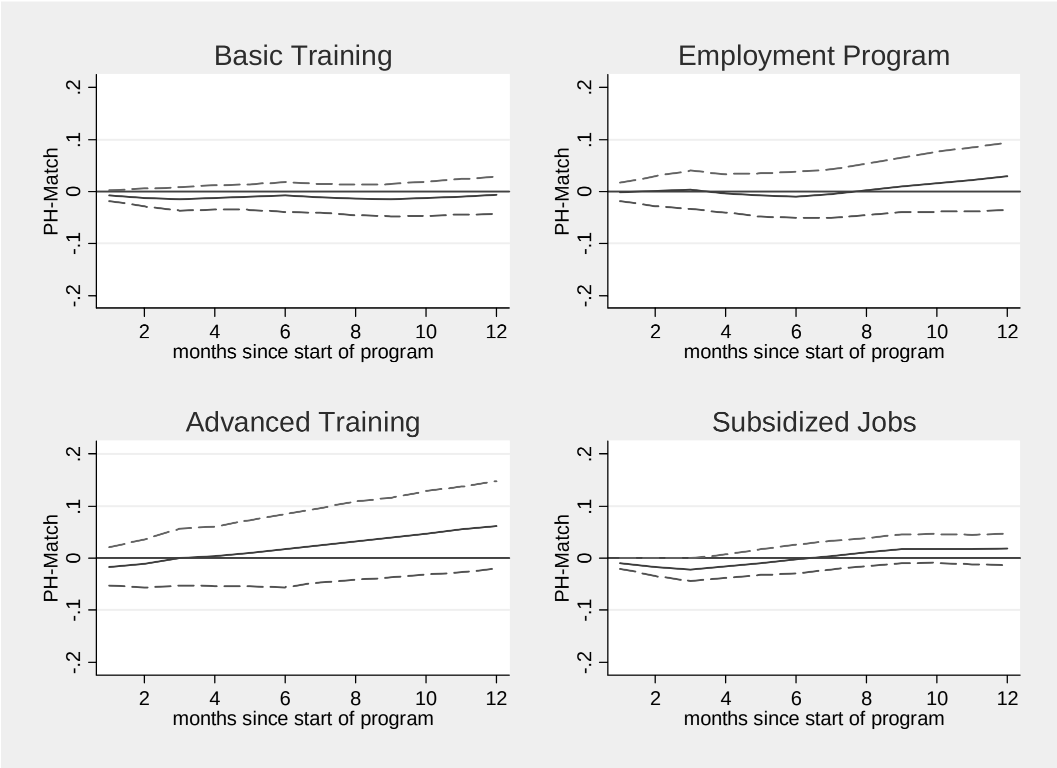


Notes: S1 is the survivor curve with treatment, S0 is the counterfactual survivor curve without treatment for the treated. Dashes lines represent 95% confidence interval based on 250 subsamples.

Source: Own calculations, based on Swiss unemployment and social security register data.

Figure 4:

Comparing the Results of Two Estimators that Use Conditional Independence
Proportional Hazard Estimator (PH) vs. Matching Estimator (Match)



Notes: PH-Match is the difference in the effect (S1-S0) according to the PH estimator and the matching estimator. Dashes lines represent 95% confidence interval based on 250 subsamples.

Source: Own calculations, based on Swiss unemployment and social security register data.

Table 1:
Descriptive Statistics of Active Labor Market Programs in Switzerland

		Unemployment Duration at Entry [Months]		ALMP-Duration [Months]		Cost per Person and Day[a)]
	# Obs.	[Mean]	[SD]	[Mean]	[SD]	[CHF]
Basic Training						
Job Training	1386	3.17	(2.57)	0.78	(1.08)	150
Language Training	573	3.63	(2.70)	2.36	(1.38)	90
Computer Training	689	3.59	(2.82)	0.96	(0.96)	170
Advanced Training						
Vocational Training	159	4.08	(3.19)	1.85	(2.04)	ca. 150
Other Training	258	3.95	(3.16)	2.18	(2.27)	ca. 150
Employment Program						
Public	415	4.56	(3.23)	4.54	(2.56)	ca. 70
Private	485	4.81	(3.46)	4.73	(2.40)	ca. 70
Subsidized Jobs	3123	2.98	(2.75)	1.55	(2.61)	-
Total	7088	3.41	(2.89)	1.83	(2.43)	

Notes: a) Maximum cost that can be refunded to program supplier.
Source: Own calculations, based on Swiss unemployment register data.

Table 2:
The Effects of Active Labor Market Programs on Transitions to Regular Jobs
Proportional Hazard Estimates

	Basic Training		Advanced Training		Employment Program		Subsidized Jobs	
A. Baseline Model								
Treatment effects (after start of program)								
0-2 months	-0.207	(-4.580)	-0.273	(-2.540)	-0.765	(-7.770)	0.014	(0.370)
3-5 months	-0.094	(-1.650)	0.144	(1.220)	0.021	(0.240)	0.170	(3.660)
6-8 months	0.076	(0.990)	0.366	(2.320)	0.339	(2.980)	0.265	(3.940)
9- months	0.365	(4.270)	0.370	(1.940)	0.340	(2.250)	0.335	(4.160)
Transition Rate to Jobs	0.079	(3.360)	0.079	(3.063)	0.079	(3.112)	0.080	(3.416)
Control Variables	Yes		Yes		Yes		Yes	
Duration Dependence	Yes		Yes		Yes		Yes	
log Likelihood	-23866.1		-19626.8		-20258.9		-25349.2	
N	15073		15073		15073		15073	
B. Constant Treatment Effect	-0.085	(-2.498)	0.010	(0.144)	-0.215	(-3.518)	0.109	(3.733)
Control Variables	Yes		Yes		Yes		Yes	
Duration Dependence	Yes		Yes		Yes		Yes	
log Likelihood	-24712.0		-20462.1		-21124.0		-26186.6	
N	15073		15073		15073		15073	

Notes: Coefficients represent effect on log hazard rate with asymptotic z-values in parentheses.
Source: Own calculations, based on Swiss unemployment insurance and social security records.

Table 3:
Effects of Active Labor Market Programs on Transitions to Regular Jobs
MMPH model that allows for endogenous censoring

	Basic Training		Advanced Training		Employment Program		Subsidized Jobs	
A. Baseline Model								
Treatment effects (after start of program)								
0-2 months	-0.306	(-6.060)	-0.392	(-3.077)	-0.912	(-7.412)	-0.074	(-1.771)
3-5 months	-0.279	(-4.391)	-0.050	(-0.381)	-0.229	(-1.887)	0.035	(0.661)
6-8 months	-0.233	(-2.696)	0.064	(0.356)	-0.035	(-0.225)	0.053	(0.688)
9- months	-0.060	(-0.572)	-0.028	(-0.122)	-0.154	(-0.750)	0.063	(0.624)
Transition Rate to Jobs								
vua	0.085	(3.032)	0.084	(2.795)	0.084	(2.842)	0.086	(3.104)
vub/vua	0.000		0.000		0.000		0.000	
Transition Rate to Program								
vpa	0.039	(1.985)	0.005	(0.810)	0.023	(1.012)	0.050	(2.077)
vpb/vpa	0.000		0.000		0.000		0.000	
Censoring Rate								
vca	0.072	(2.574)	0.110	(3.075)	0.105	(2.969)	0.059	(2.456)
vcb/vca	0.000		0.000		0.000		0.000	
Prob(vu=vua, vp=vpa, vc=vca)	0.930	(12.776)	0.931	(26.606)	0.363	(3.175)	0.926	(14.182)
Prob(vu=vua, vp=vpa, vc=vcb)	0.000		0.000		0.000		0.000	
Prob(vu=vua, vp=vpb, vc=vca)	0.000		0.000		0.576	(6.159)	0.000	
Prob(vu=vua, vp=vpb, vc=vcb)	0.000		0.000		0.000		0.000	
Prob(vu=vub, vp=vpa, vc=vca)	0.033	(0.349)	0.044	(0.459)	0.002	(0.006)	0.037	(0.458)
Prob(vu=vub, vp=vpa, vc=vcb)	0.000		0.000		0.031	(0.249)	0.000	
Prob(vu=vub, vp=vpb, vc=vca)	0.018	(0.171)	0.000		0.006	(0.075)	0.018	(0.155)
Prob(vu=vub, vp=vpb, vc=vcb)	0.019		0.025		0.022		0.019	
Control Variables	Yes		Yes		Yes		Yes	
Duration Dependence	Yes		Yes		Yes		Yes	
log Likelihood	-51101.8		-44022.0		-45562.2		-52629.9	
N	15073		15073		15073		15073	
B. Constant Treatment Effect	-0.285	(-6.759)	-0.203	(-2.353)	-0.557	(-6.890)	-0.036	(-0.975)
Control Variables	Yes		Yes		Yes		Yes	
Duration Dependence	Yes		Yes		Yes		Yes	
Unobserved Heterogeneity	Yes		Yes		Yes		Yes	
log Likelihood	-51104.6		-44025.8		-45583.4		-52632.4	
N	15073		15073		15073		15073	

Notes: Coefficients represent effect on log hazard rate with asymptotic z-values in parentheses.
Source: Own calculations, based on Swiss unemployment insurance and social security records.

Table 4:
Sensitivity Analysis: Allowing for monthly shifts in the baseline hazards
MMPH model that allows for endogenous censoring

	Basic Training		Advanced Training		Employment Program		Subsidized Jobs	
Treatment Effect								
0-2 months	-0.454	(-9.015)	-0.502	(-3.858)	-1.034	(-8.930)	-0.209	(-5.009)
3-5 months	-0.329	(-5.101)	-0.081	(-0.605)	-0.323	(-2.940)	-0.004	(-0.077)
6-8 months	-0.294	(-3.314)	0.008	(0.043)	-0.163	(-1.167)	0.004	(0.052)
9- months	-0.116	(-1.060)	-0.105	(-0.442)	-0.326	(-1.755)	-0.043	(-0.405)
Control Variables	Yes		Yes		Yes		Yes	
Duration Dependence	Yes		Yes		Yes		Yes	
Unobserved Heterogeneity	Yes		Yes		Yes		Yes	
log Likelihood	-50127.5		-43176.5		-44691.8		-51595.8	
N	15073		15073		15073		15073	

Notes: Coefficients represent effect on log hazard rate with asymptotic z-values in parentheses.
Source: Own calculations, based on Swiss unemployment insurance and social security records.

Table 5:
Sensitivity Analysis: Allowing for program start time effects
MMPH model that allows for endogenous censoring

	Basic Training		Advanced Training		Employment Program		Subsidized Jobs	
Treatment effects (after start of program)								
0-2 months	-0.157	(-2.272)	-0.307	(-1.850)	-0.754	(-4.856)	-0.001	(-0.018)
3-5 months	-0.153	(-1.994)	0.024	(0.136)	-0.098	(-0.688)	0.089	(1.529)
6-8 months	-0.136	(-1.472)	0.128	(0.616)	0.067	(0.403)	0.087	(1.109)
9- months	-0.024	(-0.226)	0.024	(0.101)	-0.116	(-0.564)	0.059	(0.585)
Treatment effect (program start time, months)	-0.057	(-3.090)	-0.026	(-0.647)	-0.050	(-1.842)	-0.032	(-2.209)
Control Variables	Yes		Yes		Yes		Yes	
Duration Dependence	Yes		Yes		Yes		Yes	
Unobserved Heterogeneity	Yes		Yes		Yes		Yes	
log Likelihood	-51096.8		-44022.0		-45560.3		-52627.4	
N	15073		15073		15073		15073	

Notes: Coefficients represent effect on log hazard rate with asymptotic z-values in parentheses.
Source: Own calculations, based on Swiss unemployment insurance and social security records.