This is a preprint of "Does the order and timing of active labor market programs matter?, Oxford Bulletin of Economics and Statistics, 72-2, 180-211, 2013, doi: 10.1111/obes.12007"

Does the order and timing of active labor market programs matter?

MICHAEL LECHNER[*] and STEPHAN WIEHLER[+]

[*]*University of St. Gallen, Swiss Institute for Empirical Economic Research (SEW-HSG), Varnbüelstr. 14, CH-9000 St. Gallen, Switzerland. (email: Michael Lechner@unisg.ch)*

[+]*University of St. Gallen, Swiss Institute for Empirical Economic Research (SEW-HSG), St. Gallen, Switzerland (email: Stephan.Wiehler@unisg.ch)*

Abstract

This paper extends the traditional focus of active labor market policy evaluation from a static comparison of participation in a program versus nonparticipation (or participation in another program) to the evaluation of the effects of program sequences, i.e. multiple participation or timing of such programs. Explicitly allowing for dynamic selection into different stages of such program sequences we analyze multiple programs, the timing of programs, and the order of programs. The analysis is based on comprehensive administrative data on the Austrian labor force. Our findings suggest that (i) active job search programs are more effective after a qualification program compared to the reverse order, that (ii) multiple participation in qualification measures dominates single participation, and that (iii) the effectiveness of several labor market programs deteriorates the later they start during an unemployment spell.

JEL classification: J68

Keywords: Active Labor market policy, matching estimation, program evaluation, panel data.

1 Introduction

For the last twenty years, a steadily growing number of active labor market policy evaluation studies have been conducted for Europe.[1] So far, the majority of studies concentrated on the impact of a single program compared to nonparticipation or to participation in other programs. In this study, we extend this focus to issues of multiple participations, as well as the timing of active labor market programs.

These issues are important for policy makers as European employment offices develop more and more sophisticated strategies for particular types of unemployed. This trend seems to be particularly pronounced in Austria. These strategies involve, for example, sequences of similar, or quite different, types of active labor market programs triggered at particular points in time in the unemployment spell. Therefore, to understand the effects of such strategies for the unemployed, it is necessary to evaluate the effects not of the single program but of the appropriate combinations of programs, for example, taking into account their specific timing. This comes with particular econometric problems, because typically the *planned* program sequence and its timing are either not known to the evaluator or not yet fully set at the beginning of the unemployment spell. Therefore, the impact of the intended sequences has to be estimated from a sample of individuals actually participating in complete sequences, instead of those assigned to the sequence. The former sample is however selective, as not all who started the sequence may reach its intended end. Instead, for reasons potentially related to the effects of the early components of the sequence, participants may drop out. This gives rise to potential dynamic selection bias and constitutes the econometric challenge tackled in this paper. To do so, we use the dynamic potential outcome model first suggested by Robins (1986) in the epidemiological literature and recently discussed by Miquel and Lechner (2001) and Lechner (2007, 2008) in the econometrics literature.

[1] For a survey on active labor market policies in Europe see Kluve et al. (2007), as well as the summaries given in Heckman, LaLonde, and Smith (1999).

So far, the conclusions from the microeconometric literature about the effects of single active labor market programs in Europe are mixed. Apart from differences at the country level, there seems to be a broad consensus about effect heterogeneity with respect to the type of the program,[2] but also with respect to program participants.[3] The majority of studies concentrate on a comparison of participation in a particular program versus nonparticipation or participation in another program, either at one point in time or in the same time span.

Recent literature has seldom addressed the effects of differential timing of labor market programs or the effects of program sequences, i.e. participating in either the same program more than once or in different programs: Carling and Richardson (2004) estimate the impact of labor market programs on subsequent unemployment in Sweden. In a duration modeling framework they find that the duration until program entry has no impact on the estimated hazard rates. Another example is Sianesi (2004) who groups Swedish labor market programs according to the length of the respective unemployment spell before the program. She finds that earlier allocation increases subsequent employment prospects. Fitzenberger and Speckesser (2007), Fitzenberger, Osikominu, and Völter (2007) and Fitzenberger and Völter (2007) adopt this approach for Germany and find similar results. However, a particular problem of that approach is that at any point in time 'starting' participants are compared with 'not yet' participants. Therefore, the implicit counterfactual changes consequently, which makes the interpretation of the estimated quantities difficult. Flores-Lagunes, Gonzalez, and Neumann (2007) estimate average causal effects of different lengths of exposure to Job Corps

[2] For job creation schemes in Switzerland see Gerfin and Lechner (2002). Similar results appear in Lechner and Wunsch (2006) and in Caliendo, Hujer, and Thomson (2004, 2006, 2008) for Germany. For wage or integration subsidies in Sweden see Sianesi (2008) and Forslund, Johannson, and Lindqvist (2004), and for Switzerland Lalive, van Ours, and Zweimüller (2002) and Gerfin, Lechner, and Steiger (2005). For business start-up programs, we refer to Carling and Gustafson (1999). For training measures, comprising formal qualification, further training of any kind, and retraining see Richardson and van den Berg (2001) for Sweden, Gerfin and Lechner (2002) for Switzerland, and Hujer, Thomsen, and Zeiss (2005) for Germany. Lechner, Miquel, and Wunsch (2004) investigate long-run effects and Lechner and Wunsch (2006) analyze business cycle effects. Winter-Ebmer and Zweimüller (1996), Hofer and Weber (2004a, b), and Lutz, Mahringer, and Pöschl (2005) investigate employment effects for different instruments of the Austrian ALMP.

[3] See Puhani (1999) and Kluve, Lehmann, and Schmidt (1999, 2004) for Poland, Friedlander, Greenberg, and Robins (1997) and Heckman, LaLonde, and Smith (1999) for the US and other western economies. Lechner, Miquel, and Wunsch (2004, 2007) for Germany, as well as a survey by Bergemann and van den Berg (2006) for gender specific differences.

(JC) training in the U.S. They suspect that the length of a program is potentially influenced by personal characteristics or by the participant's history in the labor market at any point in time. Thus, they use the idea of the 'generalized propensity score' (introduced by Imbens, 2000), i.e. the assumption that the length of the individual’s JC spell is randomly assigned, conditional on a rich set of covariates. They find a hump shaped positive relation between the program length and weekly earnings for their white and Hispanic subsamples.

As mentioned above, we use the dynamic potential outcome approach for this evaluation study. This approach requires having informative data to allow correcting for the potential dynamic selection biases. Furthermore, as the evaluation heavily relies on individuals having completed the sequences of interest, the data has to be large enough to contain a sufficiently large number of such individuals. The comprehensive administrative labor market data from Austria used in this study meets these criteria. We use it to evaluate a variety of program sequences to answer questions that are not restricted to the timing issue per se, but also relate to the issue of multiple participations in programs of a different or the same kind. Therefore, this paper is probably the first labor market evaluation study that deals with the issue of a dynamic program allocation for the case that participation in a later stage of a sequence may depend on earlier stages of the sequence and on intermediate outcomes of the latter.

Among other results, we find that earlier program allocation dominates allocation in a later stage of the unemployment spell. Furthermore, active job search programs are more effective after participating in a qualification measure, and less effective before such a program.

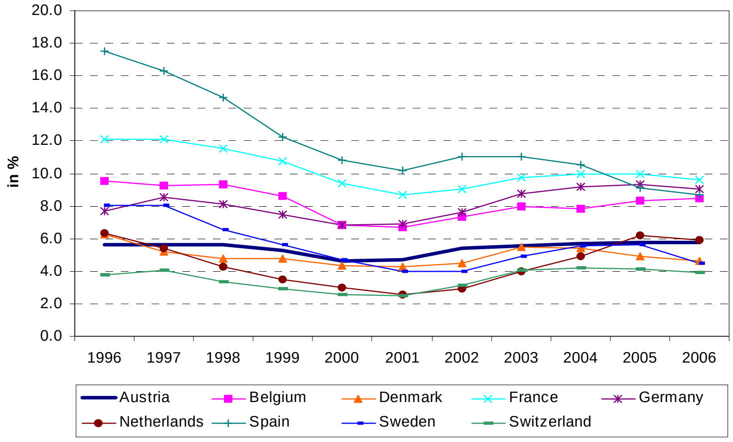
The paper is organized as follows. Section 2 briefly summarizes key features of the Austrian labor market and sketches the institutional setting. In Section 3, we introduce the data, discuss the underlying identification strategy, and provide a descriptive analysis of the participants of the various program sequences under consideration. Estimation is discussed in Section 4. The results of the estimation and a brief sensitivity analysis are provided in Section 5. Section 6 concludes. An Internet appendix with supplementary material is available on the Internet (www.sew.unisg.ch/lechner/at).

2 The Austrian labor market - development and institutions

Compared to other European countries, Austria has experienced a rather low and stable unemployment rate over the last decade after Austria's EU entry in 1995. In Figure 1 we observe a slight decrease of the unemployment rate between 1998 and 2000 with a subsequent convergence to the former level of about 5.9 percent. Only Switzerland features an unemployment rate that is lower over the entire observation period. In recent years also Denmark and Sweden undercut the Austrian unemployment rate.

The objectives for the Austrian government in fighting unemployment are legally defined in the Public Employment Service Act. In particular, they comprise (i) matching job seekers and vacancies efficiently, (ii) removing any barrier that prevents this matching, (iii) increasing the flow of information about potential matches, (iv) mitigating quantitative and qualitative differences between labor demand and supply, (v) securing existing jobs, and (vi) providing funds for the unemployed in the case of a job loss. These goals are pursued by means of a wide range of passive and active labor market policies.

Figure 1. Unemployment rates of selected European countries



Note: Data are extracted from the OECD Economic Outlook Database 2006.

Passive labor market policy in Austria covers earnings losses caused by various types of non-employment. To receive unemployment benefit (UB) payments the unemployed have to be registered at the Public Employment Service (PES), be eligible and willing to work, and have a predefined record of employment with social insurance contributions as summarized in Table 1. The standard entitlement period is 20 weeks with a replacement ratio of 55 percent of the former net wage. After UB expiry, the unemployed receive unemployment assistance (UA), which is means-tested. Ignoring any reductions due to the means-test, UA is only slightly below UB. UA has no time limit, but needs annual re-approval.

TABLE 1

Unemployment benefit duration as a function of contribution time and age

	Contribution time in weeks	Prior relevant time in months	Entitlement in weeks	Age restriction
1st application	52	24	20	no
Further applications	28	12	20	no
Prolongations	156	60	30	no
.	312	120	39	>=40
.	468	180	52	>=50
.	780	300	78	>=60

Standard replacement ratio:	55% of the previous net earnings
Family premium	Granted for dependent family members. Maximum overall replacement ratio: 80% of the previous net earnings.
Unemployment assistance:	After UB exhaustion (means tested, but no maximum entitlement period)

Table 2 shows that active labor market policies (ALMP) can be defined as belonging to one of four categories - orientation measures, training measures, subsidized employment, and others. Training measures comprise a variety of programs from active job search to different forms of qualification partially combined with on-the-job-training elements. The PES either provides programs directly or supports external providers with financial grants.

TABLE 2.

Active labor market policies - overview

Program	Abbreviation	Purpose
Orientation measures*	OM	Assessment of the individual situation and aptitude => upfront decision process for further actions.
Training		
Active job search*	AJS	Improvement of job acquisition skills, like interview training, etc.
Qualification measures*	QM	Broad class of qualification programs endowing participants with basic skills up to formal vocational degrees. Providers are connected to the PES by upfront arrangements about a fixed number of participants.
Course subsidies*	CS	Financial support for courses of external providers, i.e. not connected to the PES as for QM.
Job coaching	JC	Combination of counseling, qualification and on the job training for individuals with specific placement handicaps, like disabled persons.
Subsidized employment		
Business start-up programs	BSP	Support for participants from an initial business idea, which is worth being realized, until the actual foundation.
Socio-economic enterprise and non-profit sector projects	SEE & NPS	Subsidized employment in a quasi-realistic work environment for individuals with bad re-integration prospects.
Integration subsidies	IS	Temporary wage subsidies (up to 100%) for the first 150 days of a new employment.
Others		
Qualification for employees	QFE	Qualification measures for employed individuals.
Beneficences for labor	BFL	Collaboration with local firms in order to compensate sudden excess supply or demand of workers caused by business foundations or sudden plant closures.

Note: Measures with an asterisk are considered in the empirical analysis for reasons explained in the next section.

Due to data restrictions, we are unable to further distinguish between qualification measures and course subsidies. Subsidized employment comprises employment in a quasi-realistic work environment, business start-up measures, and integration subsidies. In addition, the PES also offers qualification measures for employed individuals that are not in the focus of this study. Overall expenditures for active labor market policies have steadily grown since 1996 and currently amount to approximately 0.6 percent of the Austrian GDP. Since the dynamic treatment evaluation approach is very data consuming, we focus on the four largest programs - orientation, qualification, active job search, and course subsidies.

Since 2000, the strategy of the Austrian PES has changed from allocating single programs to designing program sequences. In February 2001, this strategy was constituted as new component in the 'Guiding Principle of the Federal Ministry of Economics and Labor'.

As a consequence, the number of individuals attending program sequences (more than one program) within one unemployment spell increased significantly, i.e. their number almost doubled from 12,861 in 2000 to 23,560 in 2001. This strategy has remained important for the PES ever since. This institutional feature in combination with the availability of data for the entire Austrian workforce enables us to identify groups of interesting program sequences that have a sufficiently large number of participants to apply the semi-parametric dynamic evaluation framework to be discussed below.

3 Data and identification strategy

Data

As summarized in Table 3, the three data sources used in this study comprise administrative registers from the Federation of Austrian Social Insurance Institutions and the PES. The data sources are linked by a unique personal identifier. The set of variables in the last column of Table 3 are used to construct a multitude of additional variables covering the entire employment history until 1985 as well as a number of supplementary variables that must be controlled for, e.g. the remaining unemployment benefit claim at program entry or the pregnancy status for women, as is shown in Lechner and Wiehler (2007).

TABLE 3.

Administrative data used for the evaluation

Data source	Contents	Time period	Variables
Austrian Social Insurance Institutions	Times of employment (including self- and minor employment, civil servants), times of unemployment, parental leave, retirement, and employer information	1985-2005	Age, gender, nationality, gross wage, economic branch and size of the employer (in persons), regional identifier.
Austrian Public Employment Service (PES)	Times of unemployment, counseling process, times of program participation	1990-2005	Age, gender, nationality, profession, desired profession, education, family status, disability, number of job offers, type and duration of the program.

In principle information is available on a daily basis for the entire observation period. In a first step, we condense the information into a manageable form by aggregating the daily information into 2-week intervals. For reasons of tractability of the dynamic approach, a considerable part of the information will finally be aggregated over a four month period.

The parameters of interest and their identification

The approach we use in this paper is close to the dynamic causal framework suggested by Robins (1986). It is based on potential outcomes and allows for the definition of causal effects of dynamic interventions as well as systematically addressing the dynamic selection problem. By specifying the potential outcomes explicitly, and thus defining the counterfactuals of interest, the problems of the parameters proposed by Sianesi (2004) are avoided. His approach is extended in several dimensions in subsequent work[4] and frequently applied in epidemiology and biostatistics as for example in Robins, Greenland, and Hu (1999), Hernan, Brumback, Robins (2001), Keiding (2001), and Murphy (2003). A major advantage of this approach is that it allows the systematic use of time varying control variables that are partly affected by the participation in the programs under investigation. In this paper, we deviate slightly from that literature in that we rely on the identification results for parameters that are typically of interest in evaluation studies, but not necessarily so in biometrical studies (see Lechner and Miquel, 2001).

Next, using this approach we introduce the parameters of interest, discuss key identification conditions, and provide the intuition why those identifying assumptions are likely to hold true in this study. Consider a simple version of the dynamic selection problem: A group of people, unemployed in an initial period t_0, may be either sent into a program (or not) in two subsequent periods (t_1, t_2). Participation in t_1 does not exclude participation in t_2. Denoting participation in a specific period with 1 and nonparticipation with 0, all individuals

[4] For example see Robins (1989, 1997, 1999), Robins, Greenland, and Hu (1999) for discrete treatments, and Gill and Robins (2001) for continuous treatments.

will experience one of the four program sequences (0,0), (0,1), (1,0), or (1,1). The parameter of interest in this study will be the difference in outcomes for the different sequences and for a (target) population of individuals who have either been allocated to the program in t_1 or not. This effect will be denoted as the Dynamic Average Treatment Effect for a subpopulation in one sequence, DATE. Of course, simply comparing mean outcomes at $t > t_2$ does not identify any causal effect of the sequences as soon as selection into those sequences is non-random, i.e. participation status and future potential outcomes are correlated. There are many identification strategies that can be applied in the static case to resolve this issue. For example, Rubin (1974) shows how to resolve this selection problem by assuming *conditional independence* (or non-confoundedness). This assumption implies that program status and potential outcomes are independent given a set of covariates that jointly determine program participation and outcomes.[5] However, in a dynamic setting complexity increases as selection may happen every period. Controlling for selection in t_1 is no longer sufficient because selection in the second period t_2 may depend on the effects of program (non-) participation in t_1. For that reason, based on the abovementioned body of work by Robins (beginning with Robins, 1986) and his coauthors, Lechner and Miquel (2001) formulate an additional conditional independence condition for the allocation in t_2 in order to identify the dynamic treatment effects. The latter authors label it the *weak dynamic conditional independence assumption* (WDCIA). It requires that selection into the program at t_2 is independent of the potential future outcomes given the set of initial covariates, the program status the person experienced in t_1, and a set of intermediate outcomes that result from program (non-) participation in t_1. In other words, they require conditional independence in t_1, using all relevant information prior to t_1, but also conditional independence in t_2, i.e. using the program status in t_1 and all relevant information prior to t_2, which may already be influenced by the

[5] See also the extensive survey by Heckman, LaLonde, and Smith (1999). Imbens (2000) and Lechner (2001) provide identification conditions for the multiple treatment case.

first participation status. The intuition behind this approach is that if we observe all factors that drive the allocation in both t_1 and t_2 we can identify the potential outcome of a sequence for a population that did not take part in that particular sequence, but features the same distribution of characteristics at the two selection decision points in t_1 and t_2.[6] In the propensity score paradigm of Rosenbaum and Rubin (1983), Lechner and Miquel (2001) show that WDCIA can also be formulated in terms of the conditional participation probabilities in each selection step to provide identification of the respective dynamic effects.[7]

The plausibility of the WDCIA depends not only on the availability of a set of variables that drive program allocation in general, as described in Lechner and Wiehler (2007), but also on the existence and observability of relevant intermediate outcomes. Hence, apart from the entire set of variables that are deemed to be important in a static framework, our data should provide information about time varying characteristics that are presumed to be the key driving factors behind the dynamic program selection.

First, the decision of a caseworker to send the unemployed to a (further) program will certainly hinge on the labor market prospects of the unemployed individual during a sequence. The number of job offers is observed at any point in time, and certainly related to labor market prospects (although, different case workers may allocate 'their' job offers differently among the different types of unemployed). Second, the allocation will also be based on the intermediate development of the unemployed person on the labor market until program entry, e.g. intermediate employment or intermediate time outside the labor force

[6] Note that in this simple set-up the latter population is required to be defined only by the allocation in t_1. Under WDCIA identification for populations that are defined by both programs in t_1 *and* t_2 is subject to some restrictions, as described in Lechner (2007).

[7] Comparing this approach to a regression model with lagged dependent variables is tempting. However, although the static linear regression model (with the outcome variable defined as the dependent variable, and the confounders and treatment dummy defined as the independent variables) and static matching estimation are closely related if the functional form and homogeneity assumptions are correct, the relation between the dynamic regression model and the dynamic selection model based on sequential conditional independence assumptions is much vaguer. Due to the inclusion of the observed lagged outcome variable, the coefficient on the treatment dummy no longer measures the treatment effect. Furthermore, the assumptions required for consistent estimation of the regression coefficients and the treatment parameter are different (ignoring functional assumptions). See Lechner and Miquel (2001) for a detailed discussion of the relation of the dynamic treatment approach under conditional independence to classical panel data models.

(OLF). Third, the selection will most likely also be influenced by financial or other personal aspects of the unemployed person, like the remaining unemployment benefit claim, times of parental leave or even pregnancies. Since our data are very comprehensive and available on a daily basis, we are able to either observe or construct such factors at any point in time. Using 4-month windows for the time structure of the sequences, as will be explained in the next section, also guarantees that our analysis is based on sufficient variation between the different stages of a sequence.

Definition of the population and of the program sequences of interest

For the dynamic program evaluation, we impose four key requirements on the program sequences and on the underlying population that we consider. First, the identification strategy strongly hinges on the existence of a long labor market history before the entry into unemployment. Second, the follow-up period after the program sequence of interest should not be influenced by perturbing events like the possibility of (early) retirement. As a result of those arguments we concentrate on the age group of the labor force between 25 and 50 years of age. Third, the data must provide all relevant information about the selection at any stage of the program sequence. Fourth, since we employ semi-parametric estimation techniques, the number of observations in a program sequence under consideration must be sufficiently large.

Given the (time-) structure of the data and the respective information in Table 3, we focus on the first inflow of individuals from employment into unemployment or directly into a labor market program between 2000 and 2002 (labeled as the 'defining UE spell') for two reasons. We want to maximize the number of participants in each sequence and have a follow-up period of at least three years, which enables us to identify medium term effects. Table 4 summarizes the restrictions we impose on the data. We, then, are left with 345.044 individuals, on which all further steps will be based. However, so far the structure of the data is intractable for further investigations since individual unemployment careers greatly differ with respect to program participation, as well as with respect to the timing of the programs. Therefore, it is hard to find a group of people who experienced a sequence of the same order and timing on a 2-week grid.

TABLE 4.

Population reduction criterions

Criterion	Remaining individuals
All persons who become unemployed for the first time between 2000 and 2002	831,027
Unemployment without a recall guarantee (i.e. no temporary layoffs)	615,849
Age at unemployment entry between 25 and 50	395,168
Duration of the last employment > 2 months	345,044

Note: Previous employment must be longer than 2 months in order to exclude individuals in subsidized employment: With this, individuals having had a short period of unsubsidized employment (usually a couple of days) in a subsidized firm before re-entering unemployment again are excluded.

Looking at Table 5, the average time between two programs starts is approximately 4 months. A third or even a fourth program starts roughly four, respectively eight months later. Hence, we recode our data into 4-month periods denoted as trimesters from now on.[8] Of course, this is an approximation and ideally we would like to consider selection issues on a day-by-day basis. However, first of all, such a fine-time grid would render this approach hardly feasible from a practical point of view. Second, it is not necessarily required, as selection is unlikely to change so rapidly as to call for such a fine time grid.

Table 5: Average time between two consecutive program starts

Mean difference between the i th and the j th consecutive program in months			
i \ j	2	3	4
1	4.5	8.5	11.8
2	--	4.3	7.9
3	--	--	4.1

Note: Average time between two consecutive programs for the first two years after entry into the defining unemployment spell.

Having determined the population and the time structure of the data, the next step is to define program sequences suitable for our analysis. A program sequence is a chronology of trimesters in the first year after entry into the defining unemployment spell. The crucial

[8] Within each trimester we prioritize the spells that define the entire period as before, i.e. program participation followed by unemployment and employment. Whenever we find two or more programs within a trimester, we pick the last program as the defining one. Note that the defining spell does not need to cover the whole 4-month window.

requirement is that we must be able to pin down all selections along the sequence by means of a set of time invariant covariates and a set of time dependent outcomes, which were labeled as intermediate outcomes in the previous section. Sequences are denoted in brackets with the abbreviations used in Table 2. As an example, a sequence of one year with an initial participation in a qualification measure followed by two trimesters of unemployment will be labeled as (QM,UE,UE). A single (program) status in brackets, e.g. (AJS) or (UE), denotes participants in active job search or unemployment in the first trimester.

A descriptive analysis of the selection into the program sequences

Tables 6 and 7 illustrate mean characteristics for those participants per sequence that fulfill the requirements described above. The upper part contains time varying variables and the lower part time invariant variables. Table 6 shows the set of sequences that deal with the issue of program order and multiple program allocations. The first two columns contain participants who either received a qualification measure before (QM, OM) or after an orientation measure (OM, QM). In (OM, QM) we observe 20% more women. Consequently, such participants have previously spent more time on parental leave. They have lower remaining benefit claims at the entry into the second trimester, and they often have less vocational degrees. Participants in an active job search program before (AJS,QM) or after a qualification measure (QM,AJS) are similar, except that participants of (AJS,QM) have a higher fraction of pregnant participants and exhausted benefit claims before the second trimester, desire a vocational change more frequently, and are less often married at the beginning of the unemployment spell. Comparing unemployed who participated in qualification measures twice in the first two trimesters (QM, QM, UE) to those who experienced only one measure (QM,UE,UE), we observe that the latter are more likely to hold a vocational degree more often and are characterized by shorter times of previous employment and a higher fraction of exhausted benefit claims right before the third trimester. The fraction of pregnant participants right before the second trimester is lower for (QM,QM, UE). This group is also made up of a higher fraction of singles. A similar picture results from the comparison of participating in two course subsidies (CS,CS,UE) as compared to one (CS,UE,UE), both followed by employment.

TABLE 6

Mean characteristics of selected variables per sequence - order and frequency

Sequence	QM,OM	OM,QM	AJS,QM	QM,AJS	QM,QM,UE	QM,UE,UE	CS,CS,UE	CS,UE,UE
Cases	146	275	734	275	265	790	166	557
Intermediate outcomes								
Just before the second trimester of the sequence								
Intermediate time in employment (in months)	.05	.02	.01	.03	.04	.35	.04	.47
Number of transitions between UE and E	.07	.02	.03	.03	.04	.14	.07	.21
Share affected by intermediate employm.(in %)	1.1	4.1	1.4	1.5	1.9	13	4.1	19
Intermediate time 'out of labor force' (in months)	.07	.03	.03	.01	.02	.14	.03	.14
UB benefit exhausted (in %)	9	12	7	6	6	10	8	9
Cumulated job offers	1.8	1.6	1.7	1.8	1.2	2.1	3.1	3.1
Pregnant just before defining state (in %)	.69	1.1	.27	.00	.75	1.9	1.2	1.1
Just before the third trimester of the sequence								
Intermediate time in employment (in months)	--	--	--	--	.18	1.3	.32	1.3
Number of transitions between UE and E	--	--	--	--	.09	.50	.19	.60
Share affected by intermediate employment (in %)	--	--	--	--	5.7	33	13	37
Intermediate time 'out of labor force' (in months)	--	--	--	--	.06	.32	.18	.30
UB benefit exhausted (in %)	--	--	--	--	9	26	16	24
Cumulated job offers	--	--	--	--	1.5	4.2	3.3	5.7
Pregnant just before defining state (in %)	--	--	--	--	1.5	2.3	2.3	2.2
Time invariant characteristics								
Personal characteristics (in %)								
Female	59	80	64	62	66	61	58	53
Age at UE entry (in years)	37	37	37	38	36	37	37	37
Foreigner	12	14	15	12	13	14	12	12
Desires vocational change	14	15	30	25	23	27	23	20
Family status (in %)								
Single	37	33	36	32	35	28	36	36
Married	39	39	37	43	41	47	36	42
Divorced	11	15	18	14	9.4	15	14	11
Living community	6	5.5	4.1	5.5	6.4	4.2	5.2	5.2
Missing	6.9	8.2	4.4	5.8	7.6	5.2	7.5	5.2
Education (in %)								
No formal education	.00	3.6	2.5	2.6	2.6	2.5	2.3	2.2
Compulsory school	35	36	32	35	33	35	39	36
Apprenticeship	33	26	26	30	23	33	20	29
Schooling degree with vocational qualification	12	8.2	9.1	5.8	11	11	9.8	9.2
School. degr. with university entrance qualification	9.0	13	17	12	15	8.7	12	10
Academic degree	2.1	4.3	5.2	6.6	6.4	4.8	5.8	5.2
Missing	8.3	8.2	9.3	9.1	9.1	6.0	11	6.3
Last employment								
Last gross earnings in Euro per day	49	46	55	56	51	50	53	52
Duration last employment spell in months	32	33	30	34	33	24	34	27
Employment history (in months)								
Mean duration employed 5 years before UE entry	24	24	22	25	26	20	26	23
Mean duration in UE 5 years before UE entry	2.3	2.4	2	32.0	1.7	2.3	1.9	1.9
Overall time in child care	6.6	11	6.2	5.5	7.1	8.0	8.0	6.4
Fraction of entire observation period (in %)								
in unemployment	12	13	12	11	10	14	10	12
in employment	66	64	69	73	70	66	70	70
Outcomes								
Employment after end of the sequence (in%)								
12 months after UE entry	27	16	21	36	25	29	24	29
24 months after UE entry	52	51	54	62	61	55	61	61
36 months after UE entry	56	58	58	62	61	58	61	63
48 months after UE entry	52	53	60	64	63	57	59	64

Note: Time invariant characteristics are measured at the entry into the defining unemployment spell. Time dependent variables are measured until right before the spell that defines the trimester.

TABLE 7

Mean characteristics of selected variables per sequence - timing of the programs

Sequence	AJS,UE,UE	UE,AJS,UE	UE,UE,AJS	UE,QM,UE	UE,UE,QM	UE,CS,UE	UE,UE,CS
Cases	1844	2203	3387	1534	2201	828	1203
Intermediate outcomes							
Just before the 2nd trimester of the sequence							
Intermediate time in employ. (in months)	.32	.13	.32	.18	.58	.27	.69
Number of transitions between UE and E	.14	.11	.12	.14	.20	.20	.23
Share affected by interm. employm. (in %)	13	6.0	9.9	7.5	17	12	19
Interm. time 'out of labor force' (in months)	.15	.09	.16	.08	.24	.14	.25
UB benefit exhausted (in %)	10	22	12	20	11	20	10
Cumulated job offers	2.3	2.5	3.7	2.1	2.8	2.3	2.2
Pregnant just before defining state (in %)	.92	.36	.00	1.2	.05	.24	.00
Just before the 3rd trimester of the sequence							
Interm. time in employment (in months)	.88	.37	.38	.41	.72	.62	.86
Number of transitions between UE and E	.40	.22	.22	.23	.41	.34	.47
Share affected by interm. employm. (in %)	26	14	12	14	21	19	24
Interm. time 'out of labor force' (in months)	.30	.20	.20	.17	.30	.22	.32
UB benefit exhausted (in %)	27	31	77	28	66	33	62
Cumulated job offers	4.3	3.3	5.7	2.7	4.9	3.1	4.2
Pregnant just before defining state (in %)	1.8	.73	.47	1.5	.64	.60	.50
Time invariant characteristics							
Personal characteristics (in %)							
Female	49	45	45	62	55	50	48
Age at UE entry (in years)	37	38	38	37	37	37	37
Foreigner	22	24	20	15	15	17	14
Desires vocational change	34	30	27	22	23	20	19
Family status (in %)							
Single	29	29	32	32	33	31	37
Married	46	45	42	46	41	40	37
Divorced	16	15	16	12	14.2	14.5	13.4
Living community	3.7	4.7	3.8	4.5	5.1	5.1	4.5
Missing	4.6	6.6	6.0	5.7	6.0	8.2	8.2
Education (in %)							
No formal education	7.8	7.0	6.4	2.2	2.5	1.8	3.8
Compulsory school	27	29	30	37	37	35	32
Apprenticeship	36	38	40	30	31	28	31
School. degree with vocational qualification	4.9	4.5	5.0	10	8.8	7.0	8.1
School. degr. with university entrance qual.	13	7.7	8.8	11	11	11	11
Academic degree	3	4.7	3.7	3.5	4.2	5.6	6.2
Missing	8	8.5	6.3	6.5	5.9	11	7.9
Last employment							
Last gross earnings in Euro per day	52	53	51	50	52	54	54
Duration last employment spell in months	24	28	28	29	26	31	28
Employment history (in months)							
Mean duration in empl. 5 y. before UE entry	19	22	21	23	21	24	22
Mean duration in UE 5 y. before UE entry	3.2	3.0	3.8	2.4	2.8	2.1	2.5
Overall time in child care	5.2	5.4	6.7	9.5	8.4	7.1	6.5
Fraction of entire observation period (in %)							
in unemployment	15	15.6	17	13	15	12	14
in employment	67	67	65	67	66	69	66
Outcomes							
Employment after end of the sequence (in%)							
12 months after UE entry	22	22	5.0	24	4.8	26	8.5
24 months after UE entry	50	48	39	54	40	58	45
36 months after UE entry	50	51	44	57	49	61	56
48 months after UE entry	51	51	44	59	53	62	57

Note: Time invariant characteristics are measured at the entry into the defining unemployment spell. Time dependent variables are measured until right before the spell that defines the trimester.

With regard to the timing of the programs, we find some general characteristics for all sequences considered here. First, the fraction of female participants decreases as the time prior to program participation increases. The fraction of singles, in turn, increases for later program starts. Second, participants in a program at a later stage have the highest fraction of exhausted benefit claims before the start of the trimester compared to earlier participants. Over 60 percent of the participants in the last trimester have exhausted their benefit claims. Third, participants in the second trimester have the lowest time of intermediate employment until then, but also the least time out of the labor force. Participants in earlier stages of the sequence have higher pregnancy rates later on compared to the other participants. Participants in active job search in the last trimester are more likely to be skilled compared to earlier participants of this program.

Overall, participants in the different sequences differ with respect to time invariant characteristics (static selection on observables), but also with respect to intermediate outcomes at one point of time and over time. Thus, there is dynamic selection of observables into the defining states of the trimesters based on those variables.

We close this section with a remark on the choice of the sequences considered. Since we face great heterogeneity with respect to program length, we require some sequences to end with a trimester in unemployment, if population size allows. In this way, we guarantee that programs end within the sequence so that both participation groups take part in programs with similar length, which excludes the comparison of very short and long measures. Deviations from this rule are explicitly mentioned and due solely to population size restrictions. In principle, one could skip the last trimester and compare for example (QM,QM) with (QM,UE) or just (QM), again for different subpopulations. However, the degree of heterogeneity with respect to the program length and type would increase, which would make the interpretation of the estimated effects even harder, given that we are unable to distinguish certain types of qualification measures or course subsidies.

4 Estimation

General considerations

The estimator that is used in this study to compute the effects of all pair-wise comparisons of sequences of interest is the inverse probability weighting (IPW) estimator as discussed in Lechner (2007). Recall that since the participants in both sequences under consideration in any pair-wise comparison differ from the target population, it is not possible to observe the post-sequence outcomes for the latter directly in the data because of the dynamic selection issue explained in Section 3.2. Thus, the estimation strategy is to re-weigh the empirical distribution of the outcomes of those individuals who experienced an entire sequence in order to mimic the target population with respect to all of the characteristics that determine selection and outcomes at all relevant selection stages, i.e. t_1 and t_2 in the simplified world used as example in Section 3.2. Those weights are proportional to the inverse of a product of sequential probabilities of having experienced a component of the sequence conditional on what happened before (therefore, this procedure is called inverse probability weighting, IPW). Taking the sequence (0,1) as an example, this would be the conditional probability of participating in the program at t_2 given non-participation in t_1 multiplied by the conditional probability of non-participation in t_1 . Note that the underlying population for the estimation of the participation probabilities changes at every selection step. Suppose we want to re-weigh the persons in the sequence (0,1) towards a target population (1) defined in t_1 . The probability of being a member of (1) is computed on the entire population, i.e. all individuals in t_1 who are either in the program (1) or not (0).[9] The probability of participating in t_2 conditional on

[9] For this example, this automatically determines the probability of not being in the target population, i.e. being a member of (0) in t_1.

non-participation in t_1 is only computed on the subpopulation of non-participants in t_1, i.e. all individuals who are a member of (0,0) or (0,1).

In summary, the estimator is a multi-step estimator. In a first step all conditional transition probabilities are estimated, for example with probit or logit models. It is important to realize that the explanatory variables in these estimations should always be exogenous or at least predetermined, i.e. they may be influenced by past events (treatments) but not by the particular period-treatment (transition) which the particular binary choice model attempts to explain. In a second step, these transition probabilities are used to construct the weights. Lechner (2007) shows that this sequential estimator has a GMM interpretation. Thus, under standard regularity conditions of parametric GMM estimators such an estimator is $N^{1/2}$-consistent and asymptotically normally distributed for a parametric estimator of the selection probabilities.[10] The asymptotic standard errors are computed using this sequential GMM interpretation as well. For all further details, the reader is referred to Lechner (2007).

Common support

Another component of the WDCIA, not mentioned thus far, is the common support condition. Every member in one of the three populations under inspection, i.e. the two sequences and the target population, is required to have an identical counterpart with respect to the participation probability per sequence. By means of this, we create a homogeneous set of individuals who could have participated in any of the sequences with comparable propensity scores. Finally, to account for the sensitivity issue with IPW estimation concerning observations with very large weights, we trim those individuals with the largest 1 and 5 percent of weights, again, per sequence and check our results for robustness (note the relation to the concept of defining a

[10] Results of a sequential matching procedure as discussed in a previous version of Lechner (2007) are available on request. There are only small differences between the results of the matching and the IPW approach. Robins and coauthors suggest alternative estimators in various papers.

population for which the estimation can be made more precise as suggested by Crump, Hotz, Imbens, and Mitnik, 2006).

Estimation of the propensity scores

Two remarks are in order with respect to the estimation of the selection / transition probabilities: First, when there is no variation between two periods, in the sense that there is no actual transition, then it is not possible to estimate a probit. However, in this case this is not a problem because no selection can occur at this point and thus there is no correction necessary. Second, if dynamic selection is an issue, then it is most likely that participation in the past programs (states) influences the selection in future programs (states). Thus, it is important to have such time varying variables that capture these intermediate outcomes at our disposal.

In this application, the propensity scores used for the calculation of the weights are estimated by means of binary probit models for every step of the dynamic selection. For each selection step, we include a fixed set of variables that include variables capturing gender, being a foreigner, disability status, age, vocational degree, wish for a vocational change, last income, as well as summary variables of the labor market history, and regional indicators. Each specification is tested for omitted variables, stepwise, and augmented by time varying or intermediate outcomes if the respective tests indicate to do so.

TABLE 8

Results of the propensity score estimation - three selected comparisons

Variable	AJS vs. QM	(AJS,QM) if AJS	(QM,AJS) if QM	(UE,QM) vs. (UE,UE)	(UE,QM,UE) if (UE,QM)	(UE,UE,QM) if (UE,UE)	AJS vs. UE	(AJS,UE) vs. AJS	(AJS,UE,UE) if (AJS,UE)	(UE,UE) if UE	(UE,UE,AJS) if (UE,UE)
Female				+		+	+			+	-
Disabled	-	-	+	+		+	+		-	+	
Foreigner	-					-	-		-	-	
Age at UE entry	+		+	-	+	+	+	-	+	+	+
No vocational degree	-	-				-	+			+	+
University entrance qualification and academic degree	+			-		-	+		+	+	
Wish for vocational change	+			+		+	+			+	+
Month of pregnancy *							-				
Overall time in child care	-				+	+	+			+	
Last earnings			+							+	
UB claim expired*		-		-	+			+	-	+	
Mean duration											
in employment 2 years before UE entry	-			+		+	-			+	+
in unemployment 2 years before UE entry	+			-		-			+	+	+
Regional indicators											
UE rate				-		-				+	
industrial region						+				-	
Touristic region	+			+			+	+		-	
Intermediate outcomes*											
Until start of the second treatment of seq,											
intermediate time in employment (months)	A		-	-	-	+	A	+	+	-	+
Number of transitions from UE to E	A			+	+	+	A	-	+	-	+
Interm. time in 'out of labor force' (months)	A	-		-	-	+	A	+	+	-	+
UB benefit exhausted (in %)	A	+		+	-	-	A	-	+	-	-
cumulated job offers	A	-		-	-	+	A	+	-	+	+
month of pregnancy	A	-		-		-	A	+		+	
Until start of the third treatment of the seq.											
Interm. time in employment (months)	A	A	A	A	+	-	A	A		A	-
Number of transitions from UE to E	A	A	A	A	-	-	A	A	-	A	-
Interm. time in 'out of labor force' (months)	A	A	A	A	+	-	A	A	-	A	-
UB benefit exhausted (in %)	A	A	A	A		+	A	A	+	A	+
cumulated job offers	A	A	A	A	+		A	A	+	A	
month of pregnancy	A	A	A	A	+		A	A		A	-

Note: We estimate probit models for each selection step. We do not report the value of the coefficients, since they are only identified up to scale and thus not comparable between the different models. + (-) denotes that the respective variable has a positive (negative) coefficient in the index for the participation probability that is significant at the 5% level. (A) Variable is not exogenous at the beginning of the defining spell of each trimester. (*) Variables are measured right before the defining spell of the trimester. Reading example 1: For the selection into AJS compared to QM, we find a positive influence of age on the probability of participating in AJS. Reading example 2: For the selection into (AJS,QM) given the participation in AJS, we find a negative coefficient of exhausted benefit claims on the probability of getting subsequent QM. Blanks either denote that the respective coefficients are zero or that the variable is not part of the selection probit.

Time invariant covariates play a key role for the selection in the earlier phases of the sequences. Some of them, e.g. gender, disability, vocational degree, or the wish for a vocational change, appear to have a more sustainable impact also in later phases of the sequence. However, we observe subsequently that selection is driven by intermediate outcomes, like intermediate employment, intermediate times in OLF, exhausted benefit claims, cumulated job offers or pregnancies. We report estimation results of the selection probit models for selected variables for three complete comparisons (AJS, QM) vs. (QM,AJS), (UE,QM,UE) vs. (UE,UE,QM), and (AJS,UE,UE) vs. (UE,UE,AJS) in Table 8. All other probit results are available from the authors on request, but suppressed for the sake of brevity and readability.

5 Results

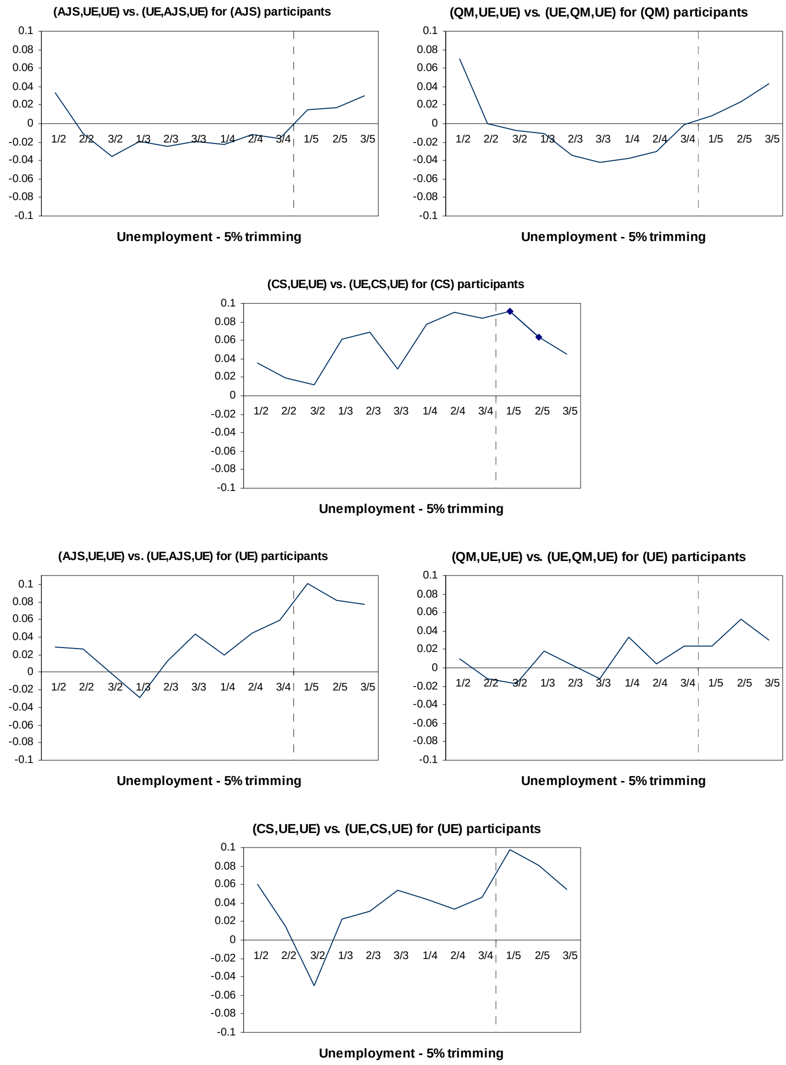
Timing of active labor market programs

This section illustrates our estimation results with regard to the timing of some active labor market programs in Austria - active job search, qualification measures, and course subsidies. We estimate the effects of timing for various outcome variables (we use indicator variables 0 and 1 for employment and unemployment and earnings in Euro) and trimming levels, but we restrict our graphical presentation to the effect on unemployment at the 5 percent trimming level for the sake of brevity. All other relevant estimation results, discussed, are summarized in the Internet appendix. The sequences under consideration and the target population are given in the first line of every graph. All outcome variables are averaged over each trimester. We measure the effects starting, from the first trimester and continuing after the end of the respective sequences. The time is measured in trimesters and years, where e.g. 1/3 denotes the first trimester in the third year after the initial entry into unemployment between 2000 and 2002. Since we face non-negligible attrition at the end of the follow-up period, we separate the last three trimesters by a dashed line and give them little weight in the interpretation of our findings. A symbol on the curves indicates that the respective effect is significant at the 5 percent level.

Figure 2 reports the effects of allocating programs in the first trimester after the initial entry into unemployment entry compared to the allocation of the same program in the second trimester. We measure this effect by comparing sequence (P, UE, UE) to sequence (UE, P, UE), as well as sequences (P, UE) to sequence (UE, P, UE), where *P* denotes one of the programs. There are two issues important to note here with respect to timing and considering periods of unemployment as part of the sequences of interest: First, as mentioned before, we require a period of unemployment after the program to ensure that program durations are comparable (thus eliminating longer programs). Furthermore, also note that in both types of comparisons we focus on individuals not working in the first two periods. Finally, using sequences of equal length has the advantage that we can compare comparable intervals, in which we require participants in both sequences not to work. However, this approach has the disadvantage of taking away some of the program effects of the (P, UE, UE) sequence as participants in that sequence are unemployed two periods after the program whereas participants in (UE, P, UE) are unemployed only one period after the program. Below we present the results for the comparisons based on sequences of equal length, and report the differences in the other comparison.

With respect to starting a program in the first or the second trimester of the unemployment spell, we do not find any persistent effects. This holds for all programs and target populations considered - active job search, qualification measures, and course subsidies.

Figure 2. Program participation in the first versus the second trimester - effect on unemployment in %



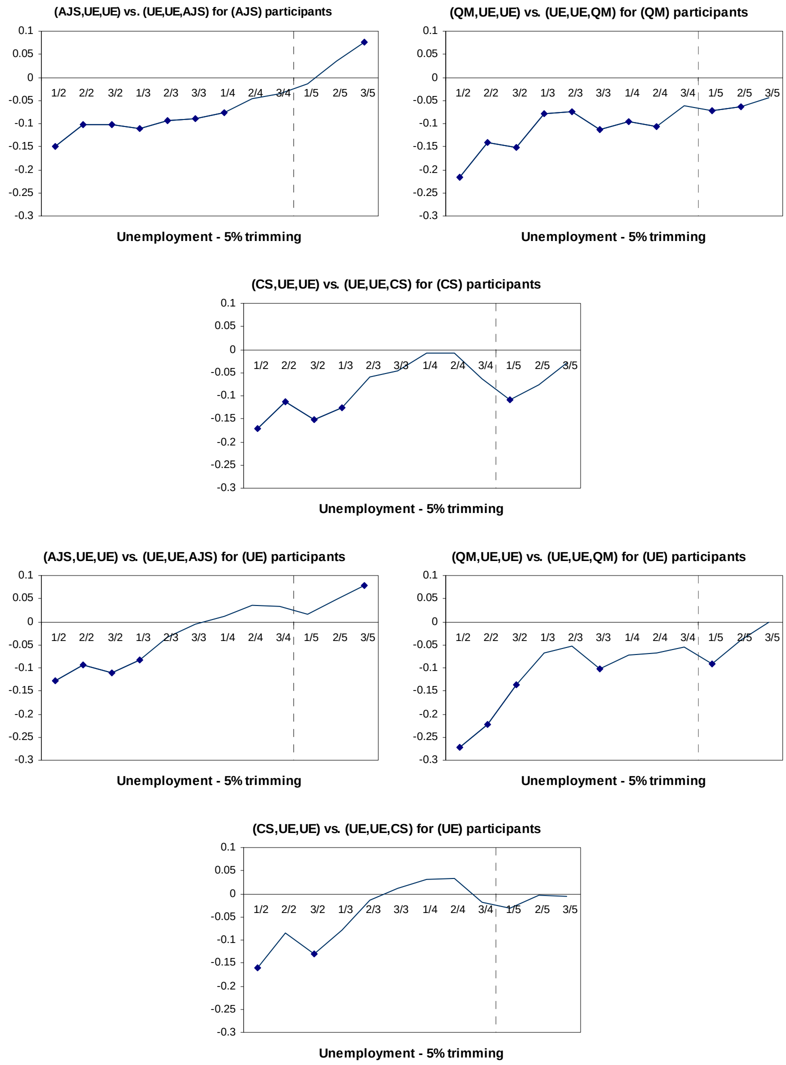
Note: AJS: active job search, QM: qualification measure, CS: course subsidy.

The results change if we consider program participation in the first or in the third trimester (Figure 3). We observe strong negative effects on unemployment right at the beginning of the follow-up period in favor of early program allocation for all programs.

However, such effects only occur because some of the individuals, who receive the program in the third trimester, are either still in the program for some time or participate in another program immediately after the first program at the beginning of the follow-up period. This so-called lock-in effect is also observed in the static treatment literature, e.g. van Ours (2004). In principle, one could eliminate this lock-in effect from the estimation by comparing, for example, (QM,UE,UE) with (UE,UE,QM,**UE**). However, conditioning also on the fourth trimester would reduce the number of participants substantially and would result in small sample problems. The lock-in effect is more pronounced for longer programs, like qualification measures, and less palpable for shorter programs, like active job search (qualification measures and course subsidies may last more than one year, whereas the duration of active job search programs is usually less than 100 days).

Ignoring the fifth year after the initial entry into unemployment, we observe that for active job search and qualification measures the reduction in unemployment due to earlier program participation does not vanish, but stabilizes at a level of about -5 to -10 percent after three years. This cannot be explained by the lock-in effect given that average program duration is considerably less than one year for all of the programs. For course subsidies we do not observe such longer-term effects.

Figure 3. Program allocation in the first versus third trimester - effect on unemployment in %

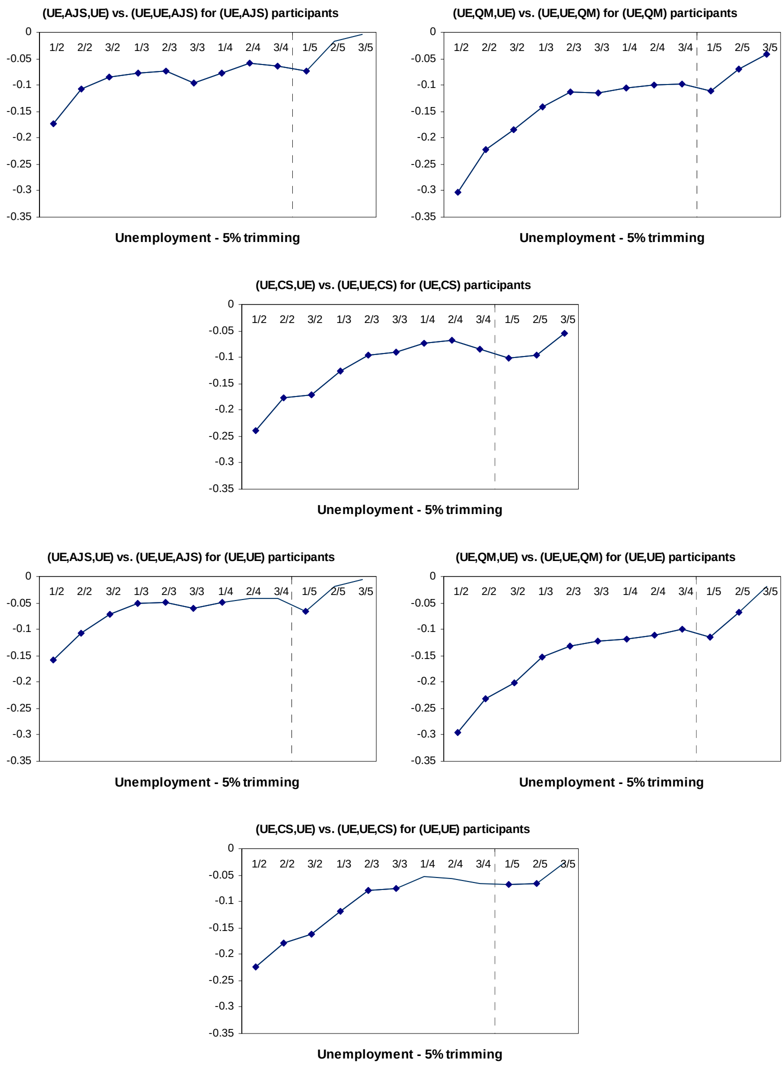


Note: AJS: active job search, QM: qualification measure, CS: course subsidy.

In Figures 2 and 3, we condition on two trimesters of unemployment after a program in the initial trimester for symmetry reasons of the sequences. Doing so, we artificially limit the composition of members in the first sequence to those individuals who fail to reenter the labor market in months 4 to 8 after the program. Hence, as mentioned before we also checked all the latter comparisons for sequences of only one UE trimester after the program, i.e. for example (AJS,UE) versus (UE,UE,AJS) or (QM,UE) versus (UE,QM,UE). As expected, we estimate larger effects in favor of early program allocation for all programs since, for example, (QM,UE) contains more individuals who find a job in the third trimester after the initial UE entry than (QM,UE,UE) by construction.

Figure 4 demonstrates a comparison of program allocations in the second trimester with an allocation in the third trimester. We again observe that earlier allocation significantly decreases subsequent unemployment after four years by 6-10 percent for all programs and target populations. The size of the effects after 3 to 4 years is similar compared to those presented in Figure 3. However, in Figure 4 there is a wider range of significant effects because in this case (see Table 7) the number of observations is greater.

Figure 4. Program allocation in the second versus third trimester - effect on unemployment in %



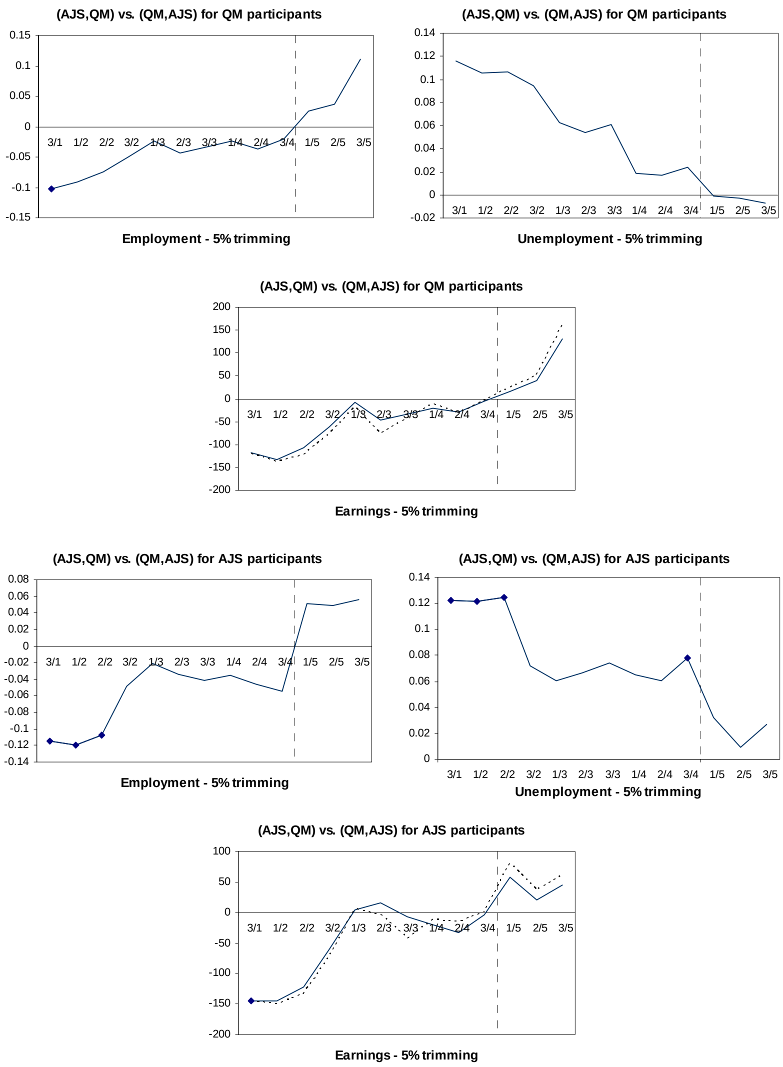
Note: AJS: active job search, QM: qualification measure, CS: course subsidy.

The effects on employment are rather similar but not identical to those on unemployment (though with the reverse sign) because we also observe times out of the labor force, i.e. neither employment nor unemployment. The effects on earnings are closely related to those on employment as earnings are only observed if a person is employed. For a more detailed treatment of earnings effects see Lechner and Melly (2007). The respective results are shown in the Internet appendix.

Order of different active labor market programs

Now, we broaden our focus from the issue of when to participate in a single program to the issue of the ordering of different kinds of programs. We focus on interesting comparisons for which the number of participants is sufficiently large enough to allow reasonable inference.

Figure 5. Order of active labor market programs - orientation and qualification measures



Note: AJS: active job search, QM: qualification measure. Dotted curve: Non-negligibly, we observe employment partially with missing earnings information. In order to include earnings in our analysis, we recode missing earnings as zero and construct a dummy variable with value one whenever we observe such a case and estimate the impact of a sequence on the latter. We then rescale the weighted sum per sequence and report the difference to assess the magnitude of the error that occurs. In general, our interpretation is not much affected by this issue. For a more elaborate treatment of earnings effects see again Lechner and Melly (2007).

Figure 5 illustrates the results of comparing participation in active job search measures before and after a qualification program - (AJS,QM) vs. (QM,AJS) - for those individuals who were either in QM or AJS in the initial trimester. For both populations we find significant lock-in effects in terms of employment at the beginning of the follow-up period for the reasons mentioned before.[11] For AJS participants in the initial trimester we find that

[11] Defining these sequences up to the third trimester, i.e. (AJS,QM,UE) versus (QM,AJS,UE), reduces the number of participants to a level that is no longer sufficient for the semi-parametric IPW estimator to deliver reliable results. Hence,

participation in active job search after a qualification measure rather than before such measure decreases unemployment four years after the initial unemployment entry significantly by 8 percent. This indicates that active job search programs should be employed when reemployment chances are potentially higher, i.e. after having participated in a qualification measure, instead of using it as a standard screening instrument per se before potential further programs. We find similar but insignificant effects for initial QM participants. Doing the same exercise for orientation measures either before or after a qualification measure does not lead to significant results.

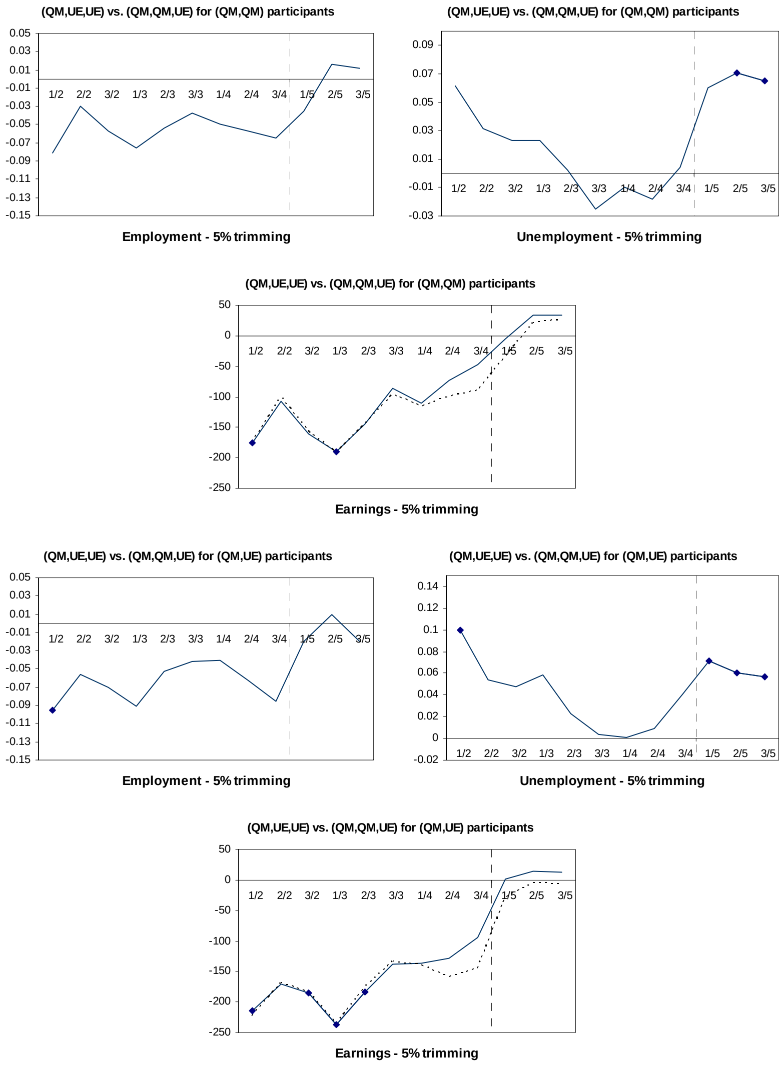
Multiple active labor market programs of the same type

As before, when we look into sequences of programs of the same type, we have to focus on program groups where such allocations can be observed in a sufficient number.[12]

we can not exclude remaining asymmetries, in the sense that some participants in (AJS,QM) have longer qualification measures than those participants in (QM,AJS) who experience programs <= 4 months by construction.

[12] Note that although the programs belong to the same type, the actual programs used in different periods, and their selection process, are most likely different.

Figure 6. Multiple allocation of active labor market programs - qualification measures



Note: UE: Unemployment, QM: qualification measure. Dotted curve: see Figure 5.

Figure 6 demonstrates the effect of a single qualification measure in the initial trimester compared to a qualification measures in the first as well as in the second trimester for both subpopulations for which such an effect is identified with the sequential randomization assumption, i.e. participants of only one or two qualification measures.[13] For both target populations, we find that double-qualification has a positive effect on employment compared to a single qualification program, though being insignificant almost everywhere. Using unemployment as the outcome variable of interest does not yield any additional insight. As far as earnings are concerned, we find that individuals, who received only one

[13] Longer programs are not counted as two single programs in two subsequent trimesters, but are excluded from this comparison.

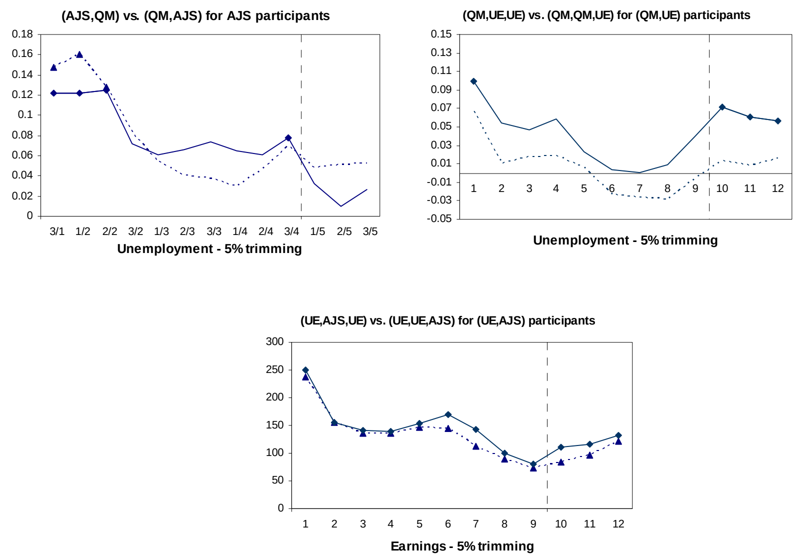
qualification, would have experienced a significant increase in gross earnings per month of about 160-250 Euros had they been treated twice. However, this effect vanishes approximately three years after the initial UE entry. This holds also for the target population that has been treated twice, but only until the first trimester of the third year after the initial UE entry. Unfortunately, our data do not allow us to distinguish between qualification measures in greater detail, which makes it difficult to derive explicit policy conclusions for this comparison. We did not find significant effects for multiple grants of course subsidies compared to single grants.

Sensitivity Checks

As an initial sensitivity check, we reduce the trimming level to 1 percent. We observe that in many cases the curves become choppier and less significant. The reason is that some persons in a sequence receiving extreme weights (who are excluded at the 5% level but included at the 1% level) basically determine the weighted sum of outcomes of the entire population of the sequence. However, the bigger picture, including all previous results, does not depend on the trimming level. Lechner (2007) also proposes a sequential matching estimator for the DATE(T). Applying this estimator in the current analysis leads to qualitatively similar conclusions.

One result of the selection analysis in Table 8 was that the allocation into the next period of a sequence is significantly determined by intermediate outcomes which are in turn influenced by earlier periods of the sequence under inspection. This constitutes evidence that it is relevant to control for such dynamics by imposing the WDCIA instead of the static CIA, which ignores selection based on intermediate outcomes. We check the relevance of the WDCIA in two additional ways: In Figure 9 we show how the results would change if we ignored the dynamic selection based on intermediate outcomes by dropping the latter and incorporating only those control variables into the selection probits that are exogenous right before the first selection. In Figure 9 the solid lines represent the effects of the previous section based on the WDCIA, whereas the dashed lines represent the results when we ignore the dynamic selection issue, and the symbols on the curves again denote significance at the 5% level. In the left panel we reconsider the comparison of using active job search before or after a qualification measure. It can be observed that ignoring the selection between different stages of a sequence leads to an overestimation of the initial lock-in effects. Furthermore, we no longer observe that using active job search after a qualification program performs better in reducing unemployment compared to the reverse order four years after the initial unemployment entry.

Figure 9. Checking the relevance of the WDCIA - selection models based only on control variables that are exogenous before the first period

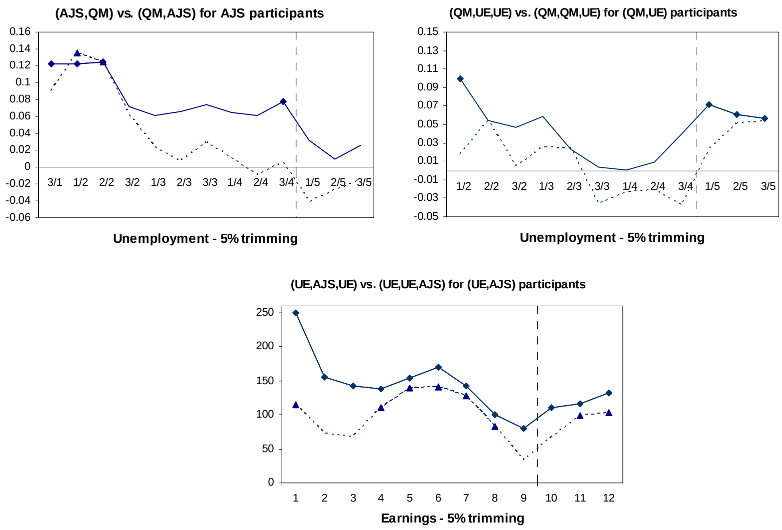


Note: UE: Unemployment, AJS: active job search, QM: qualification measure. We denote the effects that were estimated without the intermediate outcomes by a dotted line. Symbols on the curves denote again significance at the 5% level. Solid lines represent the effects of the previous section, i.e. based on the WDCIA, whereas dashed lines represent the results that ignore the dynamic selection issue.

In the second panel, we see that the estimated effects on unemployment of participating in a qualification program once rather than twice turn out to be 3-4 percent lower when intermediate outcomes are ignored. The same holds for the right panel where the estimated effects for the timing of active job search result in up to 10-35 Euros less per month.

The results in Figure 10 are based on the assumption that the intermediate outcomes are exogenous and are thus valid control variables for the selection into the sequences. In this case, the dynamic selection problem collapses to a static problem of entering one of the sequences. The sequences can then be compared directly ignoring the dynamics of the selection process by using a static evaluation approach with all intermediate outcomes as control variables. In all three panels we find that the size of the effects changes once we impose that assumption. Under the stronger selection assumption the order of active job search and qualification measures has no impact on employment in the longer-term. Furthermore, an earlier timing of active job search seems to have a much smaller effect on earnings as compared to the estimator valid under the weaker sequential independence assumption.

Figure 10. Checking the relevance of the WDCIA - selection models based on all control variables that are observed until the last selection step



Note: UE: Unemployment, AJS: active job search, QM: qualification measure. We denote the effects that were estimated without the intermediate outcomes by a dotted line. Symbols on the curves denote again significance at the 5% level. Solid lines represent the effects of the previous section, i.e. based on the WDCIA, whereas dashed lines represent the results that ignore the dynamic selection issue.

Finally, we find that accounting for the dynamic selection into different phases of a sequence matters for two reasons. First, it helps to get a better understanding of the dynamic selection within a sequence. Second, it has a clear impact on the magnitude of the estimated dynamic treatment effects under consideration.

6 Concluding remarks

This paper deviates from the traditional (static) concept of evaluating single program participation versus nonparticipation or participation in another program towards the concept of evaluating program sequences, i.e. considering sequences of (multiple) programs partially interrupted by times of unemployment. We explicitly allow for sequential allocation into the different stages of the sequence based on previous states and intermediate outcomes. We employ an inverse probability weighting estimator and use large administrative data of the Austrian labor force.

Our findings can be divided into two parts. Regarding the timing, we observe that program participation in active job search, qualification measures, and course subsidies in the first and second trimester after the initial unemployment entry performs better in terms of reducing unemployment compared to participating in the third trimester. This result can be reconciled with Sianesi (2004) for Sweden and Fitzenberger and Speckesser (2007) for Germany who also find that earlier allocation into programs is beneficial for the unemployed.

Analyzing sequences of programs of a different or the same type, we find that active job search programs are more effective after a qualification measure compared to the reverse order. This is intuitive given the purpose of active job search. We find no evidence for effects of the order of orientation and qualification measures. In addition, we find that a sequence of two qualification measures in the first and second trimester performs better in terms of earnings four years after the initial unemployment entry compared to only one qualification program in the initial trimester. Our sensitivity checks illustrate that the estimated effects of some sequences change considerably once we ignore the dynamic selection issue.

The unique information content and size of our data allowed us to gain a first set of answers to questions of how programs should be timed and how often or in what order they should be allocated. Extensions to a wider range of programs or sequences with other program constellations are interesting. They are left for future research as even in our data the respective subsamples become too small for a reliable semiparametric analysis.

References

Austrian Institute of Economic Research (2005): 'Evaluierung der Österreichischen Arbeitsmarktförderungen 2000-2003', Wien.

Bergemann, A., and G. van den Berg (2006): 'Active Labor Market Policy Effects for Women in Europe: A Survey,' Discussion Paper 2365, Institute for the Study of Labor (IZA).

Caliendo, M., R. Hujer, and S. Thomsen (2004): 'New Evidence on the Effects of Job Creation Schemes in Germany - A Matching Approach with Threefold Heterogeneity,' *Research in Economics*, 58, 257-302.

Caliendo, M., R. Hujer, and S. Thomsen (2006): 'Sectoral Heterogeneity in the Employment of Job Creation Schemes in Germany,' *Jahrbücher für Nationalökonomie und Statistik*, Vol. 226, pp. 139-179.

Caliendo, M., R. Hujer, and S. Thomsen (2008): 'Identifying Effect Heterogeneity to Improve the Efficiency of Job Creation Schemes in Germany,' *Applied Economics*, Vol. 40, pp. 1101-1122.

Carling, K., and L. Gustafson (1999): 'Self-Employment Grants vs. Subsidized Employment: Is There a Difference in the Re-Employment Risk?,' Working Paper, Institute for Labor Market Policy Evaluation (IFAU).

Carling, K., and K. Richardson (2004): 'The Relative Efficiency of Labor Market Programs: Swedish Experience from the 1990s,' *Labor Economics*, Vol. 11, pp. 335-354.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik (2006): 'Moving the Goalposts: Addressing Limited Overlap in the Estimation of Average Treatment Effects by Changing the Estimand,' NBER Technical Working Papers, 0330.

Fitzenberger, B. and S. Speckesser (2007): 'Employment Effects of the Provision of Specific Professional Skills and Techniques in Germany', *Empirical Economics*, Vol. 32, pp. 529-573.

Fitzenberger, B., A. Osikominu, and R. Völter (2007): 'Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany', unpublished manuscript, University of Frankfurt.

Fitzenberger, B. and R. Völter (2007): 'Long-Run Effects of Training Programs for the Unemployed in East Germany', *Labour Economics*, Vol. 14, pp. 730-755.

Flores-Lagunes A., A. Gonzalez, and T. Neumann (2007): 'Estimating the Effects of Length of Exposure to a Training Program: The Case of Job Corps', Discussion Paper 2846, Institute for the Study of Labor (IZA).

Forslund, A., P. Johannson, and L. Lindqvist (2004): 'Employment Subsidy - A Fast Lane from Unemployment to Work?,' Working Paper 2004:18, Institute for Labor Market Policy Evaluation (IFAU).

Friedlander, D., D. Greenberg, and P. Robins (1997): 'Evaluating Government Training Programs for the Economically Disadvantaged,' *Journal of Economic Literature*, Vol. 35, pp. 1809-1855.

Gerfin, M., and M. Lechner (2002): 'A Microeconometric Evaluation of the Swiss Active Labor Market Policy,' *The Economic Journal*, Vol. 112, pp. 854-893.

Gerfin, M., M. Lechner, and H. Steiger (2005): 'Does Subsidized Temporary Employment Get the Unemployed Back to Work? An Econometric Analysis of two Different Schemes,' *Labor Economics*, Vol. 12, pp. 807-835.

Gill, R., and J. Robins (2001): 'Causal Inference for Complex Longitudinal Data: The Continuous Case', *The Annals of Statistics*, 2001, pp. 1-27.

Hernan, M, B. Brumback, and J. Robins (2001): 'Marginal Structural Models to Estimate the Joint Causal Effect of Nonrandomized Treatments,' *Journal of the American Statistical Association*, Vol. 65, pp. 331-366.

Heckman, J., R. LaLonde, and J. Smith (1999): *The Economics and Econometrics of Active Labor Market Programs*, vol. III A of *Handbook of Economics*. North-Holland, Amsterdam.

Hofer, H., and A. Weber (2004a): 'Are Job Search Programs a Promising Tool? A Microeconometric Evaluation for Austria,' Discussion Paper 1075, Institute for the Study of Labor (IZA).

Hofer, H., and A. Weber (2004b): 'Employment Effects of Early Interventions in Job Search Programs,' Discussion Paper 1076, Institute for the Study of Labor (IZA).

Hujer, R., S. Thomsen, and C. Zeiss (2005): 'The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany,' Working paper, Goethe University Frankfurt.

Imbens, G. (2000): 'The Role of the Propensity Score in Estimating Dose-Response Functions,' *Biometrica*, Vol. 87, pp. 706-710.

Keiding, N. (1999): 'Event History Analysis and Inference From Observational Epidemiology,' *Statistics in Medicine*, Vol. 18, pp. 2353-2363.

Kluve, J., H. Lehmann, and C. Schmidt (1999): 'Active Labor Market Policies in Poland: Human Capital Enhancement, Stigmatization, or Benefit Churning?,' *Journal of Comparative Economics*, Vol. 27, pp. 61-89.

Kluve, J., D. Card, M. Fertig, M. Gora, L. Jacobi, P. Jensen, R. Leetmaa, L. Nima, E. Patacchini, S. Schaffner, C. Schmidt, B. v.d. Klaauw, and A. Weber (2007): 'Active Labour Market Policies in Europe', Springer, Heidelberg.

Lalive, R., J. van Ours, and J. Zweimüller (2002): 'The Impact of Active Labor Market Programs on the Duration of Unemployment,' Working Paper ISSN 1424-0459, Institute for Empirical Research in Economics (IEW).

Lechner, M. (2001): 'Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption', in: M. Lechner and F. Pfeiffer (eds.), *Econometric Evaluation of Active Labour Market Policies*, pp. 43-58 Heidelberg: Physica.

Lechner, M. (2008): 'Matching Estimation of Dynamic Treatment Models: Some Practical Issues', in D. Millimet, J. Smith, and E. Vytlacil (eds.), *Advances in Econometrics, Volume 21, Modelling and Evaluating Treatment Effects in Econometrics*.

Lechner, M. (2007): 'Sequential Causal Models for the Evaluation of Labor Market Programs', forthcoming in the *Journal of Business & Economic Statistics*.

Lechner, M., and B. Melly (2007): 'Earnings Effects of Training Programs', Working Paper, University of St. Gallen.

Lechner, M., and R. Miquel (2001): 'A Potential Outcome Approach to Dynamic Programme Evaluation – Part I: Identification', Discussion paper 2001-07, Department of Economics, University of St. Gallen; revised 2005.

Lechner, M., R. Miquel, and C. Wunsch (2005): 'Long Run Effects of Public Sector Sponsored Training in West Germany,' CEPR Discussion Paper 4851, Centre for Economic Policy Research.

Lechner, M. and C. Wunsch (2006): 'Are Training Programmes More Effective When Unemployment is High?,' Discussion Paper 2006-23, University of St. Gallen.

Lechner, M. and C. Wunsch (2007): 'The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany After Unification,' *German Economic Review*, Vol. 8, pp. 468–507.

Lechner, M., and S. Wiehler (2007): 'Kids or Courses? Gender Differences in the Effects of Active Labor Market Policies,' Discussion Paper 2740, Institute for the Study of Labor (IZA).

Lechner, M., and C. Wunsch (2006): 'Active Labor Market Policy in East Germany: Waiting for the Economy to Take off,' Discussion Paper 2006-24, University of St. Gallen.

Lutz, H., H. Mahringer, and A. Pöschl (2005): 'Schwerpunkt 1: Verhinderung und Bekämpfung der Langzeitsarbeitslosigkeit und der Jugendarbeitslosigkeit,' WIFO, IHS, L&R: *Evaluierung Europäischer Sozialfonds 2000-2006*, Ziel 3: Österreich. Wien.

Murphy, S. (2003): 'Optimal Dynamic Treatment Regimes,' *Journal of the Royal Statistical Society*, Vol. 65, pp. 331-366.

Puhani, P. (1999): Evaluating Active Labor Market Policies - Empirical Evidence for Poland during Transition, ZEW Economic Studies 5, Heidelberg: Physica.

Richardson, K., and G. van den Berg (2001): 'The Effect of Vocational Employment Training on the Individual Transition Rate from Unemployment to Work,' *Swedish Economic Policy Review*, Vol. 8, pp. 175-213.

Robins, J. (1986): 'A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Periods -Application to Control of the Healthy Worker Survivor Effect,' *Mathematical Modelling*, Vol. 7, pp. 1393-1512.

Robins, J. (1989): 'The Analysis of Randomized and Nonrandomized AIDS Treatment Trials Using a New Approach to Causal Inference in Longitudinal Studies', in Sechrest, L., Freeman, H., Mulley, A. (Eds.), *Health Service Research Methodology: A Focus on Aids*, pp. 113-159, Washington, D.C., Public Health Service, National Centre for Health Services Research.

Robins, J. (1997): *Causal Inference from Complex Longitudinal Data. Latent Variable Modelling and Applications to Causality*, in Berkane, M. (Ed.), Lecture Notes in Statistics (120), Springer, New York, pp. 69-117.

Robins, J. (1999): 'Association, Causation, and Marginal Structural Models', *Synthese,* Vol. 121, pp. 151-179.

Robins, J. S. Greenland, and F. Hu (1999): 'Estimation of the Causal Effect of a Time-varying Exposure on the Marginal Mean of a Repeated Binary Outcome', *Journal of the American Statistical Association*, Vol. 94, pp. 687-700.

Rosenbaum, P., and D. Rubin (1983): 'The Central Role of the Propensity Score in Observational Studies for Causal Effects,' *Biometrica,* Vol. 70, pp. 41-55.

Rubin, D. (1974): 'Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies,' *Journal of Educational Psychology*, Vol. 66, pp. 668-701.

Sianesi, B. (2008): 'Differential Effects of Swedish Active Labor Market Programmes for Unemployed Adults During the 1990s,' *Labour Economics,* Vol. 15, pp. 370-399,.

Sianesi, B. (2004): 'An Evaluation of the Swedish System of Active Labour Market Programs in the 1990s', *Review of Economics and Statistics,* Vol. 86, pp. 133-155.

Van Ours, J. (2004): 'The Locking-in Effect of Subsidized Jobs', *Journal of Comparative Economics,* Vol. 32, pp. 37-52.

Winter-Ebmer, R., and J. Zweimüller (1996): 'Manpower Training Programmes and Employment Stability,' *Economica*, Vol. 63, pp. 113-130.