published in Journal of the Royal Statistical Society,
Series A, Statistics in Society, 165, 59-82, 2002

Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods

Michael Lechner*

University of St. Gallen

Swiss Institute for International Economics and Applied Economic Research (SIAW)

First draft: June 2000

Revised: September 2001

Date this version has been printed: 21 June 2005

Forthcoming in the *Journal of the Royal Statistical Society – Series A*

Keywords: Multiple programmes, programme evaluation, treatment effects, balancing score, matching, sensitivity analysis.

JEL classification: C10, C50, J60, J68.

Addresses for correspondence

Michael Lechner

Swiss Institute for International Economics and Applied Economic Research (SIAW)

University of St. Gallen

Dufourstr. 48, CH-9000 St. Gallen, Switzerland

Michael.Lechner@unisg.ch, www.siaw.unisg.ch/lechner

---

* I am also affiliated to CEPR, London, IZA, Bonn, and ZEW, Mannheim. Paper has been prepared for the international RSS conference on THE EVALUATION OF ECONOMIC AND SOCIAL POLICIES in London, July, 4, 2000. Financial support from the Swiss National Science Foundation (NFP 12-53735.18) is gratefully acknowledged. The data are a subsample from a database prepared for the evaluation of the Swiss active labour market policy together with Michael Gerfin. I am grateful to the State Secretariat of Economic Affairs of the Swiss Government (seco) and the Bundesamt für Sozialversicherung for providing the data and to Michael Gerfin for his help in preparing them. I also thank Markus Frölich for comments and for providing me with an improved version of the Gauss code for the multinomial probit model. Helpful comments from the editor and two anonymous referees of this journal are gratefully acknowledged. All remaining errors are my own.

Abstract

Recently several studies have analysed active labour market policies using a recently proposed matching estimator for multiple programmes. Since there is only very limited practical experience with this estimator, this paper checks its sensitivity with respect to issues that are of practical importance in this kind of evaluation study. The estimator turns out to be fairly robust to several features that concern its implementation. Furthermore, the paper demonstrates that the matching approach per se is no *magic bullet* solving all problems of evaluation studies, but that its success depends critically on the information available in the data. Finally, a comparison with a bootstrap distribution provides some justification for using a simplified approximation of the distribution of the estimator that ignores its sequential nature.

1 Introduction

Many European countries use substantial active labour market policies (ALMP) in order to bring Europe's notoriously high levels of unemployment back to some sort of socially acceptable level by increasing the employability of the unemployed. These policies consist typically of a variety of subprogrammes, such as employment programmes, training, and wage subsidies, among others.

Recent evaluation studies surveyed for example by Fay (1996) and Heckman, LaLonde, and Smith (1999) do not appear to develop any consensus whether these programmes are effective for their participants. Quite to the contrary, many studies raise serious doubts. However, it could be argued that the policy implications of many of these studies were limited because their econometric framework was not ideally suited to the problem, and because the available data used were typically far from being ideal as well.

Recently the Swiss government encouraged several groups of researchers to evaluate the Swiss active labour market policies using administrative data from the unemployment registers and the pension system. Among those studies were also two econometric studies by Lalive, van Ours, and Zweimüller (2000), and Gerfin and Lechner (2000). The former paper uses a structural econometric modelling approach based on modelling the duration of unemployment, whereas the latter uses an extension of an essentially nonparametric pseudo-experimental matching approach to multiple treatments proposed and discussed by Lechner (2000b, 2001a). The fact that these studies use different (more or less explicit) identification strategies points to the issue that for every evaluation study there is the crucial question of which identification strategies and estimation method would be suitable for the specific situation. Angrist and Krueger (1999), Heckman and Robb (1986), and Heckman, LaLonde, and Smith (1999) provide an excellent overview of available identification and resulting estimation strategies.

Of course the choice of an identification strategy is strongly linked to the type of data available about the selection process into the programmes. Gerfin and Lechner (2000) argue that they observe the major variables influencing selection as well as outcomes, so that the assumption that labour market outcomes and selection are independent conditional on these observables (the conditional independence assumption, CIA) is plausible. Being able to use CIA for identification in combination with having a large data set has implications for the choice of a suitable estimator. Desirable properties of an estimator in this situation are it should avoid almost any other assumption than CIA, such as functional form assumptions for specific conditional expectations of the variables of interest. In particular the estimator of choice should avoid restricting the effects of the programmes to be the same in specific subpopulations because there is substantial a priori evidence that those programmes could have very different effects for different individuals (effect heterogeneity). Finally, this ideal estimator has to take account of the very different programmes that make up the Swiss ALMP (programme heterogeneity). To be able to convince policy makers about the merits of the results of any evaluation, the estimator needs to be based on a general concept, that could easily be communicated to noneconometricians.

An estimator that is nonparametric in nature, allows for effect as well as programme heterogeneity, and that is based on a statistical concept easy to communicate, is the recently suggested matching estimator for heterogeneous programmes. The general idea of matching is to construct an artificial comparison group. The average labour market outcomes of this group is compared to the average labour market outcomes of the group of programme participants. When CIA is valid, this estimator is consistent when the selected comparison group and the group in the specific programme have the same distribution of observable factors determining jointly labour market outcomes and participation. Matching for binary comparisons has recently been discussed in the literature and applied to various evaluation problems by Angrist (1998), Dehejia and Wahba (1999), Heckman, Ichimura, Smith, and Todd (1997), Lechner (1999, 2000a), and Smith and Todd (2000), among others. The standard matching approach that considers only two states (for example in the programme vs. not in the programme) has been extended by Imbens (2000) and Lechner (2001a) to allow for multiple programmes.

The results by Gerfin and Lechner (2000) indicate considerable heterogeneity with respect to the effects of different programmes. They find substantial positive employment effects for one particular programme that is a unique feature of the Swiss ALMP. It consists of a wage subsidy for temporary jobs in the regular labour market that would otherwise not be taken up by the unemployed. They also find large negative effects for traditional employment programmes operated in sheltered labour markets. For training courses the results are mixed.

There is only very limited practical experience with these kind of matching estimators for multiple programmes (to the best of our knowledge, the only other applications of this specific approach are Brodaty, Crepon, and Fougère, 2001, Dorsett, 2001, Frölich, Heshmati, and Lechner, 2000, Larsson, 2000, and Lechner, 2000b). In particular Lechner (2000b) discusses issues relevant for the implementation of the estimator. Here we cover several other points that could be potentially responsible for the results obtained by Gerfin and Lechner (2000). It is of particular interest whether the stark differences of the effects for the two different types of subsidised employment are robust in these respects. In addition, the sensitivity of the results to the amount of information included in the estimation will be addressed. Obviously, robustness of the results should not be expected in that exercise.

The plan of the paper is as follows: The next section summarises the results for multiple treatments obtained in Lechner (2001a) and describes the proposed estimator. Section 3 briefly discusses several aspects of the application. Section 4 presents the results of the baseline specification. Section 5 discusses the sensitivity of results by considering several deviations from the baseline specification. Section 6 concludes.

2 Econometric framework for the estimation of the causal effects

2.1 Notation and definition of causal effects

2.1.1 Notation

The prototypical model of the microeconometric evaluation literature is the following: An individual can choose between two states (causes). The potential participant in a programme gets an hypothetical outcome (e.g. earnings) in both states. This model is known as the Roy (1951) - Rubin (1974) model of potential outcomes and causal effects (see Holland, 1986, for an extensive discussion of concepts of causality in statistics, econometrics, and other fields).

Consider the outcomes of $(M+1)$ different mutually exclusive states denoted by $\{Y^0, Y^1, ..., Y^M\}$. Following that literature the different *states* are called *treatments*. It is assumed that each individual receives only one of the treatments. Therefore, for any individual, only one component of $\{Y^0, Y^1, ..., Y^M\}$ can be observed in the data. The remaining M outcomes are counterfactuals. Participation in a particular treatment m is indicated by the variable $S \in \{0,1,...M\}$.

2.1.2 Pair-wise effects

Assuming that the typical assumptions of the Rubin model are fulfilled (see Holland, 1986, or Rubin, 1974, for example), equation (1) defines pair-wise average treatment effects of treatments *m* and *l* for the participants in treatment *m*:

$$\theta_0^{m,l} = E(Y^m - Y^l | S = m) = E(Y^m | S = m) - E(Y^l | S = m). \qquad (1)$$

$\theta_0^{m,l}$ denotes the expected effect for an individual randomly drawn from the population of participants in treatment *m*. Note that if participants in treatments *m* and *l* differ in a way that is related to the distribution of attributes (or exogenous confounding variables) *X*, and if the treatment effects vary with *X*, then $\theta_0^{m,l} \neq -\theta_0^{l,m}$, i.e. the treatment effects on the treated are not symmetric.

2.2 Identification

2.2.1 The conditional independence assumption

The framework set-up above clarifies that the average causal effect is generally not identified. Therefore, this lack of identification has to be overcome by plausible, untestable assumptions. Their plausibility depends on the problem analysed and the data available. The papers by Angrist and Krueger (1999), Heckman and Robb (1986), and Heckman, LaLonde, and Smith (1999) provide an excellent overview about available identification strategies in different situations.

Imbens (2000) and Lechner (2001a) consider identification under the conditional independence assumption (CIA) in the model with multiple treatments. CIA defined to be valid in a subspace of the attribute space is formalised in expression (2):

$$Y^0, Y^1, \ldots, Y^M \amalg S | X = x, \forall x \in \chi. \qquad (2)$$

This assumption requires the researcher to observe all characteristics that jointly influence the outcomes as well as the selection into the treatments. In that sense, CIA may be called a 'data hungry' identification strategy. Note that CIA is not the minimal identifying assumption, because all what is needed to identify mean effects is conditional mean independence. However, CIA has the virtue of making the latter valid for all transformations of the outcome variables. Furthermore, in most empirical studies it would be difficult to argue why conditional mean independence should hold and CIA might nevertheless be violated.

In addition to independence it is required that all individuals in that subspace actually have the possibility to participate in all states (i.e. $0 < P(S = m \mid X = x)$, $\forall m = 0,...,M$, $\forall x \in \chi$). This condition is called the common-support condition and is extensively discussed in Lechner (2001b). Note that for any pair-wise comparison it is sufficient that for all values of X for which the treated have positive marginal probability, there could be comparison observations as well.

Lechner (2001a) shows that CIA identifies the effects defined in equation (1). Indeed, Gerfin and Lechner (2000) argue that their data is so rich, that it seems plausible to assume that all important factors that jointly influence labour market outcomes and the process selecting people into the different states can be observed. Therefore, CIA is the identifying assumption of choice. In Section 4 we elaborate on the actual identification in this application.

2.2.2 Reducing the dimension using balancing scores

In principle the basic ingredients of the final estimator would be estimators of expressions like $E(Y^l \mid X, S = l)$, because CIA implies that $E(Y^l \mid S = m) = E_X[E(Y^l \mid X, S = l) \mid S = m]$. However, nonparametric estimators could be problematic, because of the potentially high dimensional X and the resulting so called *curse of dimensionality*. For two treatments however Rosenbaum and Rubin (1983) show that conditioning the outcome variable on X is not necessary, but it is sufficient to condition on a scalar function of *X*, namely the participation probability conditional on the attributes (this is the so-called balancing score property of the propensity score). For the case of multiple treatments Lechner (2001a) shows that some modified versions of the balancing score properties hold in this more general setting as well.

Denote the marginal probability of treatment *j* conditional on *X* as $P(S = j \mid X = x) = P^j(x)$. Lechner (2001a) shows that the following result holds for the effect of treatment *m* compared to treatment *l* on the participants in treatment *m*:

$$\theta_0^{m,l} = E(Y^m \mid S = m) - \underset{P^{l|ml}(X)}{E}[E(Y^l \mid P^{l|ml}(X), S = l) \mid S = m]. \tag{3}$$

$$P^{l|ml}(x) = P^{l|ml}(S = l \mid S \in \{l, m\}, X = x) = \frac{P^l(x)}{P^l(x) + P^m(x)}.$$

If the respective probabilities $P^{l|ml}(x)$ are known or if a consistent estimator is available, the dimension of the estimation problem is reduced to one. If $P^{l|ml}(x)$ is modelled directly, no information from subsamples other than the ones containing participants in *m* and *l* is needed for the identification and estimation of $\theta_0^{m,l}$ and $\theta_0^{l,m}$. Thus, we are basically back in the binary treatment framework.

In many evaluation studies considering multiple exclusive programmes it seems natural to jointly specify the choice of a particular treatment from all or a subset of available options. $P^{l|ml}(x)$ could then be computed from that model. In this case, consistent estimates of all marginal choice probabilities $[P^0(X), ..., P^M(X)]$ can be obtained. Hence, it may be attractive to condition jointly on $P^l(X)$ and $P^m(X)$ instead of $P^{l|ml}(X)$. $\theta_0^{m,l}$ is identified in this case as well, because $P^l(X)$ together with $P^m(X)$ is 'finer' than $P^{l|ml}(X)$:

$$E[P^{l|ml}(X) \mid P^l(X), P^m(X)] = E[\frac{P^l(X)}{P^l(X)+P^m(X)} \mid P^l(X), P^m(X)] = P^{l|ml}(X). \tag{4}$$

2.3 A matching estimator

Given the choice probabilities, or a consistent estimate of them, the terms appearing in equations (3) can be estimated by any parametric, semiparametric, or nonparametric regression method. One of the popular choices of estimators in a binary framework is matching (for recent examples see Angrist, 1998, Dehejia and Wahba, 1999, Heckman, Ichimura, Smith, and Todd, 1997, Lechner, 1999, 2000a, Smith and Todd, 2000). The idea of matching on balancing scores is to estimate $E(Y^l \mid S=m)$ by forming a comparison group of selected participants in *l*, that have the same distribution of the balancing score (here $P^{l|ml}(X)$ or [$P^l(X)$, $P^m(X)$]) as the group of participants in *m*. By virtue of the property of being a balancing score, the distribution of *X* will also be balanced in the two samples. The estimator of $E(Y^l \mid S=m)$ is the mean outcome in that selected comparison group. Typically, the variances are computed as the sum of empirical variances in the two groups (ignoring the way the groups have been formed). Compared to non-parametric regression estimates, a major advantage of matching is its simplicity and its intuitive appeal. Advantages compared to parametric approaches are its robustness to the functional form of the conditional expectations (w.r.t. $E(Y^l \mid X, S=l)$) and that it leaves the individual causal effect completely unrestricted and hence allows arbitrary heterogeneity of the effects in the population. Lechner (2000b, 2001a) proposes and compares different matching estimators that are analogous to the rather simple matching algorithms used in the literature on binary treatments. The exact matching protocol that is used for the application is based on $[P^l(X),\ P^m(X)]$ and detailed in Table 1.

--- Table 1 about here --

Several comments are in order. Step 2 ensures that we estimate only effects in regions of the attribute space where two observations from two treatments can be observed having a similar participation probability (common support requirement). Otherwise the estimator will give biased results (see Heckman, Ichimura, Smith, and Todd, 1998).

A second remark with respect to the matching algorithm concerns the use of the same comparison observation repeatedly in forming the comparison group (*matching with replacement*). This modification of the 'standard' estimator is necessary for this estimator to be applicable at all when the number of participants in treatment *m* is larger than in the comparison treatment *l*. Since the role of *m* and *l* could be reversed in this framework, this will always be the case when the number of participants is not equal in all treatments. This procedure has the potential problem that a few observations may be heavily used although other very similar observations are available. This may result in a substantial and unnecessary inflation of the variance. Therefore, the potential occurrence of this problem should be monitored.

A third remark concerns the appearance of the variables $\tilde{x}$ in Step 3 b). This subset of conditioning variables already appears in the score. The motivation for including them also explicitly in the matching is that they are potentially highly correlated with the outcome variables (but not influenced by them) as well as with selection. Therefore, it seems to be particularly important to obtain very good matches with respect to these variables even in smaller samples. Although by the virtue of the balancing score property, including them as additional matching variable is not necessary asymptotically because they are already included in the score. Note that including them in the score as well as additional matching variables amounts to increasing the weight of these variables, suspected to be critically important, when forming the matches.

3 Application

The application in this paper is based on the evaluation study of the different programmes of the Swiss ALMP by Gerfin and Lechner (2000). They focus on the individual success in the labour market that is due to these programmes. The Swiss government made available a very informative and large data base consisting of administrative records from the unemployment insurance system as well as from the social security system. It covers the population of unemployed persons in December 1997. Gerfin and Lechner (2000) claim that in these data all major factors that jointly influence both the selection into the various programmes as well as employment outcomes are observed.

Let us very briefly reconsider their main line of argument to establish identification. First note that the decision to participate in a programme is made by the case worker according to his impressions obtained mainly from the monthly interviews of the unemployed. In order to evaluate this 'subjective impression' the law requires that programmes must be necessary and adequate to improve individual employment chances. Although the final decision about participation is always made by the case worker (or somebody the case worker has to report to), the unemployed may also try to influence this decision during the conversations that takes place in these interviews. Furthermore, although the law is enacted at the federal level, the 26 Swiss cantons exercise considerable autonomy in interpreting and implementing the rules specified in this law. To summarise, it does not appear to be possible to exactly state how an individual participation decision is made, but it should be possible to specify the information set on which this decision is based. Luckily, all the information obtained by and available to the case worker is stored in a centralised data base to which we have access to and which is described below. To that data coming from the unemployment registrars we add information on the last ten years of labour market history coming from the pension system. We suspect that labour market experience influences the individual preferences considerably, although it might be argued that the relevant part for selection and outcome is already contained in the data base coming from the unemployment registrar. In the following the database, the sample, as well as the programmes are briefly described.

The data from the unemployment registrars cover the period January 1996 to March 1999 for all persons who were registered as unemployed on December 31, 1997. These data provide very detailed information about the unemployment history, ALMP participation and personal characteristics. The pension system data cover 1988-1997 for a random subsample of about 25'000 observations. The exact variables used in this study can be found in Appendix WWW that can be downloaded from the internet (www.siaw.unisg.ch/lechner/l_jrss_a). They cover sociodemographics (age, gender, marital status, native language, nationality, type of work permit, language skills); region (town/village and labour office); subjective valuations of the case worker (qualifications, chances of finding a job), sanctions imposed by the placement office; previous job and desired job (occupation, sector, position, earnings, full / part-time); a short history of labour market status on a daily basis, and the employment status and earnings on a monthly basis for the last ten years. Gerfin and Lechner (2000) apply a series of sample selection rules to the data. The most important ones are to consider only individuals unemployed on Dec 31, 1997 with an unemployment spell of less than one year who have not participated in any major programme in 1997 and are aged between 25 and 55.

The active labour market programmes (ALMP) can be grouped into three broad categories: a) training courses, b) employment programmes (EP), and c) temporary employment with wage subsidy (TEMPORARY WAGE SUBSIDY, TEMP). The former two groups are fairly standard for a European ALMP encompassing a variety of different programmes. The last type of programme is rather unique, however. The difference between b) and c) is that employment programmes take place outside the “regular” labour market (see below). By contrast TEMP refers to a regular job.

In this study we focus on a subset of programmes, namely COMPUTER COURSES (COC), EP, and TEMP (and NONPARTICIPATION, NONP) (the first participation in programme with a duration of more than two weeks, starting after January 1, 1998, decides the assignment to the respective group; any programme participation later on is treated as being the effect of the first programme). The effects for these programmes were the most interesting ones found in Gerfin and Lechner (2000). Note that the validity of CIA allows us to analyse the effects of these programmes on the subsample of nonparticipants and participants in the respective programmes thus avoiding any selectivity bias problem that arises from ignoring individuals in other programmes not considered here. The reduction of the sample has the important advantage for this paper that computation times are considerably reduced.

A problem concerns the group of nonparticipants. For this group important time varying variables like 'unemployment duration prior to the programme' are not defined. To make meaningful comparisons to those unemployed entering a programme, in the baseline estimate an approach suggested in Lechner (1999) is used: for each nonparticipant a hypothetical programme starting date from the sample distribution of starting dates is drawn. Persons with a simulated starting date later than their actual exit date from unemployment are excluded from the data set. Later on in Section 5.1 other ways to handle this problem will be presented. Note that deleting nonparticipants could potentially bias the results of the effects of the programmes on nonparticipants, because it changes the distribution of nonparticipants by deleting systematically individuals with higher unemployment probability. However, this has no implication for effects defined for any of the populations of participants, which are typically those of policy interest.

-- Table 2 about here ---

Table 2 shows the number of observations as well as some descriptive statistics for subsamples composed of nonparticipants as well as participants in the three programme groups considered. The mean duration of the programme is just one month for computer courses and almost 150 days for employment programmes. The table shows that important variables like qualification, nationality and duration of unemployment also vary substantially. The final column indicates that the employment rate at the last day in our data varies considerably between 26% and 48%. Of course, this is not indicative of programme success because the composition of different groups of participants differ substantially with respect to variables influencing future employment, so that we expect differences for these different groups of unemployed even when they would not have participated in any programme.

4 Results for the baseline scenario

4.1 Selection into the programmes

The baseline scenario basically reproduces the results obtained by Gerfin and Lechner (2000) for the sample used here. The first step is an estimation of the conditional probabilities of ending up in each of the four states. The full set of the estimation results of a multinomial probit model (MNP) using simulated maximum likelihood with the GHK simulator and 200 draws for each observation and choice equation (e.g. Börsch-Supan and Hajivassiliou, 1993) and Geweke, Keane, and Runkle, 1994) can be found in Appendix WWW.

The variables that are used in the MNP are selected by a preliminary specification search based on binary probits (each relative to the reference category NONP) and score tests against omitted variables. The final specification contains a varying number of mainly discrete variables that cover groups of attributes related to personal characteristics, valuations of individual skill and chances in the labour market as assessed by the placement office, previous and desired future occupations, and information related to the current and previous unemployment spell, and past employment and earnings. Note that variables that are only related to selection and not to the potential outcomes need not be included for consistent estimation.

In practice, some restrictions on the covariance matrix of the errors terms of the MNP need to be imposed, because not all elements of it are identified and to avoid excessive numerical instability. Here all correlations of the error terms with the error term of the reference category are restricted to zero. The covariance matrix is not estimated directly, but the corresponding Cholesky factors are used.

The results are in fact very similar to those obtained by Gerfin and Lechner (2000), hence the reader is referred to that paper for the detailed interpretation. Here it is sufficient to note that there is considerable heterogeneity with respect to the selection probabilities. Again we find that better 'risks' (in terms of unemployment risk) are more likely to be in COC, whereas 'bad risks' are more like to be observed in EP.

-- Table 3 about here ---

Table 3 shows descriptive statistics of the estimated probabilities that are the basis for matching. In particular there is a large negative correlation between the probabilities of TEMP and EP with NONP.

4.2 Matching

The number of deleted observations due to the common support requirement across different subsamples is given in Table 4. The criterion used is that all estimated marginal probabilities are larger than the smallest maximum of the corresponding probability in any sample. The reverse must hold for minima. The share of observations lost varies between subsamples, but they are very small, never exceeding 3% in this paper. In contrast, Gerfin and Lechner (2000) find a reduction of more than 14% due to so–called LANGUAGE COURSES whose participants are very different from the rest of the unemployed. These courses are omitted from the current analysis. For a detailed discussion of issues related to the common-support problem, see Lechner (2001b).

-- Table 4 about here --

-- Table 5 about here --

Since one-to-one matching is with replacement, there is the possibility that an observation may be used many time thus inflating the variance. Table 5 presents the share of the weights of the 10% of observations that have been used most (i.e. 10% of those matched comparisons with the largest weights are matched to *number-in-table %* of the treated; this concentration ratio must of course be larger than 10% which corresponds to the case when every comparison observation is used only once). Given the limited experience with this approach the respective numbers appear to be in the usual range. It is however obvious that the smaller the sample the smaller the diversity of the probabilities so that the same observations are used more frequently.

Checking the match quality with respect to several variables including the probabilities used for matching shows that the matched comparison samples are very similar to the treated samples.

4.3 Effects

The measure for the success of the programme is employment in the regular labour market at any given time after the start of the programme. Hence the outcome variable is binary. The time on the programme is not considered to be regular employment. Due to data limitations the potential period of observing programme effects cannot be longer than 15 months, because the latest observation dates from March, 31, 1999. In that sense the analysis will be restricted to the short run effects of the ALMP.

-- Table 6 about here ---

Table 6 displays the mean effects of the programmes on their respective participants one year after the individual programme participation starts. The entries in the main diagonal show the employment rates in the four groups in percentage points. The programme effects are off the main diagonals (to ease reading and writing in most cases NONP is called a programme). A positive number indicates that the effect of the programme shown in the row compared to the programme appearing in the column is an on average higher rate of employment for those who participate in the programme given in the row (example: the mean effect of TEMP compared to COC is 8.0 %-points of additional employment for participants in TEMP).

The results for the respective participants in the programmes (upper part of Table 5) indicate that TEMP is superior to almost all other programmes. The mean gain compared to the other programmes is between about 6% and 16%-points. In particular TEMP is the only programme that dominates NONP. On the other hand EP have negative effects. COC are somewhat intermediate in general, but they do look fairly bad for their participants.

Figure 1 shows the dynamics of the effects by pinning down their development over time after the start of the programme. It presents the pair-wise effects for all programmes and their respective participants. A value larger than zero indicates that participation in the programme would increase employment chances compared to being allocated to the respective other programme.

Considering the relative positions of the respective curves, the line for NONP reveals the expected profile (Fig. 1a – 1c): In the beginning it is positive and increasing, but then it starts to decline as participants leave their respective programmes and increase their job search activities. Overall the findings set out in Table 6 are confirmed: TEMP dominates. EP are dominated by NONP and TEMP. For those participating in EP there is no significant difference compared to participating in COC. For the participants in COC there is a small positive initial effect compared to EP. This effect is probably due to the fact that COC are much shorter than EP.

-- Figure 1 about here --

5 Sensitivity analysis

There is only very limited practical experience with these kind of matching estimators for multiple programmes. In particular Lechner (2000b) discusses several topics relevant for the implementation of the estimator. Here, these considerations are extended to cover several other issues that could be potentially responsible for the results obtained in the study by Gerfin and Lechner (2000). In addition to these the sensitivity of the results with respect to the amount of information included in the estimation will be addressed.

The various topics are structured in the following way. In subsection 5.1 some fundamental specification problems directly related to identification are discussed. Subsection 5.2 is devoted to issues that could be considered as being of a technical nature relating to the implementation of the estimator and to obtaining valid inference.

5.1 Fundamental issues

5.1.1 Unknown start date of counterfactual programme

Most active labour market policies have the feature that individuals will enter the various programmes at different times. Here, entries into the first programme are stretched over a period of 13 months (January, 2, 1998, to January, 31, 1999), however, about half of the entries are observed in the first quarter of 1998. The information about the start of the programme plays a role in two respects: First, it is used directly in the first step of the estimation (MNP) and to compute several variables, like unemployment duration before the programme, that are assumed to be important in affecting programme participation and outcomes. Thus they are important to achieve identification. Second, the effect of the programmes is measured after their start.

Note that there is a decision to be made on how to use or generate start dates. This decision obviously concerns nonparticipants, but in principle it is also relevant for participants of other programmes. The question is always '*when would the comparison person have started the programme?*'. In the absence of any better hypothesis for participants, it appears natural to assume that the start date is actually independent of the specific programme the person is allocated to. In this case the observed start date could be used as counterfactual start date for all other programmes. If the start date is also independent of the characteristics of the individual, a natural choice for the participants is a random draw from the distribution of the observed start dates of all participants. For the binary treatment framework, other alternatives are discussed in Lechner (1999) that are applicable here as well. However, mainly due to their additional complexity they are less attractive in a multi-programme evaluation that is more computer intensive than a binary one. Of course this procedure needs another adjustment for the case when the simulated start date is in contradiction to the institutional arrangements (here, an individual needs to be unemployed to enter a programme). In the baseline scenario this approach is used and 'contradictory' nonparticipants, i.e. those with on average shorter unemployment spells (37% of all nonparticipants), are deleted from the sample.

Although in specific applications the assumption of random start dates could be plausible, it is probably more plausible to assume that start dates could be predicted by the variables influencing outcomes and selection (as long as they do not depend on the start date). Again in this case, using the observed start dates for the participants seems to be the best choice. For the nonparticipants start dates should be drawn from the conditional distribution of start dates given the covariates. As a sensitivity check, the log of start dates (the earliest day is 2, the latest is 391) are regressed on covariates, with start date dependent covariates substituted by proxies (unemployment duration is approximated by unemployed duration end of 1997, for example). To simulate the start date a log normal distribution is assumed for the start day based on a linear specification of its conditional mean (taken from the regression). It turns out that start dates can to some extent be predicted using these covariates, although an R^2 of 5% shows the limited amount of useful information contained in the covariates with respect to the timing of the programmes. The number of observations deleted reduces to 28%. In another check, this approach is used on a subsample of participants that enter the programme only in the first quarter of 1998, thus making the start date distribution more homogenous. In this case the reduction of the sample of participants resulted in a loss of 50% of the participants. Only 12% of the nonparticipants are deleted.

In order to avoid flooding the reader with numbers Table 7 shows only the effects of NONPARTICIPATION for nonparticipants, because they should be most sensitive to these changes in the specification. It appears that despite the considerable reduction in sample size in the final specification the sensitivity to these variations of the specification is small. This is confirmed by checking the dynamic patterns (Figure 2). No substantial differences can be discovered, other than an increased variance due to the smaller sampler.

--- Table 7 about here --

--- Figure 2 about here --

5.1.2 Available information

The data used for the empirical study is exceptional in the sense that it contains rich information about the current unemployment spell and previous employment histories. It is argued that such informative data is necessary to make the CIA a valid identifying assumption. In this subsection we check how sensitive the results are with respect to that information. In addition to the baseline specification, the following specifications are considered (note that each specification is less informative than the previous one):

- *No long term history*: No information from the pension system about the last ten years.
- *No duration of current unemployment spell.*
- *No subjective information:* No subjective information on employment chances as given by the case worker.
- *No information on current unemployment spell.*
- *No information about previous employment, skills, and occupation.*
- *No regional information.*
- *Only age, gender and marital status (no information on language and citizenship).*
- *No information (unadjusted differences).*

Table 8 shows the effects for different specifications for one particular set of pair-wise effects, namely the effects of COC for participants in such courses. A priori one would expect to see the most substantial changes here, because the participants appear to be clearly a positive selection in terms of unemployment risk, in particular as compared to participants in EP.

--- Table 8 about here --

The results are indeed sensitive to shrinking the information set. Let us first consider the effects of COC compared to EP. Initially there is a small negative effect of COC that is however insignificant. By removing information about the individual work related characteristics the effect increases monotonically up to a level of 15%. It is only the removal of the regional information that does not change the estimates (conditional on the information available in the previous step). So, obviously, participants of COC and EP have different chances in the labour market and that any estimate of the effects needs to take account of these differences in order to avoid substantial biases in the estimated effects.

For the comparisons of COC with NONP and with TEMP – both programmes have less pronounced differences in the attributes of its participants compared to COC – the changes can be substantial but they are not necessarily monotonous, suggesting that in this case it is not necessarily 'better' to control for more variables than for 'fewer'.

The results from Table 8 are confirmed by considering the dynamics in Figures 3. Although the patterns in all comparisons change, it is again the comparison between COC and EP that exhibits the largest effect.

Finally, one remark is in order with respect to the information contained in the subjective valuation of the labour offices. The changes in the estimate suggest that this information may indeed be valuable in uncovering characteristics that would otherwise be left undetected (of course this statement is conditional on the information set used here).

--- Figure 3 about here --

5.2 Technical issues

5.2.1 Issues related to the first step of the estimation

The specification of the conditional probabilities could also have an influence on the results. A first decision to make is whether the conditional participation probabilities should be estimated for each combination of states separately as binary choices, or whether the process should be modelled simultaneously with a discrete choice model including all relevant states. The former has the advantage of being a more flexible specification, whereas the latter is much easier to monitor and to interpret. Lechner (2000b) devoted considerable attention to this problem and found that for a very similar application, nothing was gained by going the more flexible route of modelling the binary choices separately. When using a multinomial discrete choice model a flexible version appears to be desirable. However, the computational costs may be substantial. The multinomial probit model (MNP) estimated by simulated maximum likelihood seems to be an attractive compromise, because it is sufficiently fast to compute but does not impose the so-called independence of irrelevant alternatives assumption (IIA), which the multinomial logit model does.

To check the sensitivity of the results with respect to the specification of the covariance matrix of the errors terms appearing in the MNP choice equations, the covariance between the error terms of COC and all other alternatives are set to zero. Furthermore, the sensitivity of the results with respect to the number of simulations used in the GHK simulator is checked by computing the results for just 2 draws as well as 800 draws, whereas the baseline specification is based on 200 draws per choice equation and observation.

--- Table 9 about here --

Again, since the results for COC could be expected to be most sensitive to those changes, they are presented in Table 9 and Figure 4. From the results concerning the number of draws these issues do not appear to matter at all, because all changes are of an order of magnitude of less than half a standard deviation of the estimator. The sensitivity with respect to the covariance structure is however larger (more than one standard deviation in the comparison to NONP). On the one hand this finding suggests that using a discrete choice model that relies on a more restrictive specification, like the multinomial logit model, could lead to biases. One the other hand, there could be an argument for avoiding multinomial models altogether and using (many) binary models instead.

--- Figure 4 about here --

5.2.2 The common support requirement

The conditional independence assumption implies that the decision to participate can be considered as random conditional on the covariates. To be nontrivial 'randomness' requires that for a given vector of covariates there is a positive probability of participating in every programme. The first step to ensure this requirement is satisfied in an application is to consider only individuals who – according to the institutional settings – could in principle participate in the programmes under consideration. In the current study this refers to the requirement that individuals have to be unemployed in Dec, 31, 1997 (in addition to some other requirements, see Gerfin and Lechner, 2000). As a property of a MNP the estimated conditional probabilities for all individuals are strictly bounded away from zero. However, we may find (extreme) values of the covariates that generate conditional probabilities for participants in one programme that cannot be found for participants in other programmes. Hence, there is no way to estimate the effect for this (extreme) group with the sample at hand. At this point there are two ways to proceed. the obvious one is to ignore this problem by refering to asymptotics: although the probabilities of being of observed in a particular state with such covariates may be very small, eventually (which means with some other random sample) there will be such an observation and matching will be satisfactory. Of course, with the data at hand there will be a (finite sample) bias if the potential outcomes vary with the probabilities, because these (extreme) cases lead to bad matches. The second option is to ensure that the distributions of the balancing scores overlap by removing extreme cases. The drawback here is that the definition of the treatment effects are changed in the sense they are now mean effects for a narrower population defined by the overlap in the support.

Table 4 already showed the loss of observations when restricting the sample by considering the smallest maximum and the largest minimum in the subsamples as joint bounds for the common support. The overall loss of observations is rather small. One could argue that the density in the tail of the implied distributions is still very thin, because there could be a substantial distance for example from the smallest maximum to the second smallest element of that probability in this specific subsample. Therefore, to check the sensitivity a stricter requirement is imposed, where the maximum and the minimum is substituted by the tenth largest and tenth smallest observation. The suspicion that the density may be thin seems to be justified, because the number of observations lost due to that more restrictive requirement increase from about 1-3% (see Table 4) to 16% for NONP, 14% for COC, 15% for EP, and 19% for TEMP. Because TEMP seems to be most affected by these changes, Table 9 as well Figure 5 show the effects for this programme.

--- Table 9b about here --

-- Figure 5 about here --

When the common support condition is not enforced, the major change is that the positive effect with respect to NONP is reduced and is no longer significant at the 5% level, which changes indeed one important policy conclusion. Another change concerns the increased effect in comparison to EP. However, this increase by 1%- point is less than half a standard deviation of the estimator and hence it is not substantial. To summarise, these results tend to suggest the importance of removing extreme observations. Since matching is with replacement and the samples are large, additional trimming of thin tails seems not to be necessary, at least in this application.

The paper by Lechner (2001b) suggests another way in addition to the conventional removal of observations. The idea entertained there is that although the original effect of interest is not identified without common support, the information available may nevertheless be used to obtain sharp bounds in cases when the expectation of the outcome variable is finite with known lower and upper limits.

5.2.3 Asymptotic distribution

This study so far conducted inference based on the presumption that the estimators have an asymptotic normal distribution derived from the difference of two weighted means of independent observations. This approximation however ignores the fact that the comparison group is formed by matching using an estimated balancing score based on a simulation of start dates for nonparticipants. Furthermore, estimated probabilities are used for the data driven reduction of the sample to ensure the common support criterion. So far no asymptotic theory taking account of these features of the estimator has been developed. One way to check the accuracy of this approximation for the current study is to compare the approximation to an inference based on bootstrapping. Since each estimation is fairly expensive in terms of computation time, the bootstrap is based on only 400 bootstrap samples. For each estimation a new sample of the same size is drawn with replacement and all steps of the estimation, including simulation of start dates and the enforcement of common support, are performed on the simulated sample.

Table 10 compares several estimates obtained from the bootstrap samples with those obtained from the approximation. Quite arbitrarily the results are only given for TEMPORARY WAGE SUBSIDY. However, the other results are similar. The table displays the results for the mean, the standard deviation, and some quantiles commonly used in inference. Since the more extreme quantiles could be subject to considerable simulation error due to the small number of bootstrap replications, the 25% and the 75% quantile are given as well. In addition Fishers test for normality based on the skewness and kurtosis of the distribution of the effect across the bootstrap samples is shown. In turns out that the results based on the approximation and those based on the bootstrap are fairly similar. There is probably a slight underestimation of the variability of the estimates by the approximation.

--- Table 10 about here --

Figure 6 presents the corresponding dynamics for all treatments. An effect is only displayed if the upper and lower bound of the 95% empirical bootstrap interval have the same sign. It is very hard to spot any difference between Figure 1 and the bootstrap results. Thus the baseline results are again confirmed. Given the computer intensity of the bootstrap for large samples, the approximation retains a considerable attractiveness. However, the usual pace of development in computer technology may change that statement in the (near?) future.

--- Figure 6 about here --

6 Conclusion

The study by Gerfin and Lechner (2000) analysed the Swiss active labour market policy using a newly proposed matching estimator for multiple programmes. The study is based on rich data, so that conditioning on the information available in that data, selection to the various programmes and the outcome variables are likely mutually independent. Furthermore, sample sizes are comparatively large.

In such a situation the matching estimator in its multiple programme version is an attractive choice. It has the advantage that it is basically nonparametric or at least semiparametric so that very few additional assumptions are necessary at the estimation stage of the analysis. Furthermore, it allows the effect to vary across individuals and programmes in an unrestricted way. Finally, the principles underlying this estimator are fairly easy to communicate to non-statistical consumers of evaluation studies.

There is only a very limited practical experience with these kind of matching estimators for multiple programmes. In this paper the sensitivity of this estimator with respect to some features that are of importance in empirical studies is checked. It turns out that the estimator is fairly robust to several issues that concern its implementation. The only exception to some extent is the specification of the probability model used to predict the various participation probabilities that form the basis for matching. The comparison with a bootstrap distribution provides some justification for the common use of a simplified approximation of the distribution of the matching estimator that ignores several issues relating to its sequential nature.

The paper also demonstrates that the matching approach per se is no *magic bullet* solving all problems of evaluation studies, but that its success depends critically on the information available in the data, i.e. whether using the conditional independence assumption for identification is plausible. Given the obvious insight that the performance of this estimator depends on the information available, any discussion whether this or any other estimator is the 'best' estimator for evaluation studies in general is obviously misguided.

Although matching cannot solve all potential problems of an evaluation study, if identification can be achieved by rich data and sufficient institutional knowledge about the selection process, then it is the opinion of the author that some version of matching is clearly the estimator of choice. However if the conditional independence assumption is not plausible, then there is no a priori reason why matching should be any better than any other evaluation estimator. In this case the researcher has to decide whether to collect more data, or to find another plausible identifying assumption.

References

Angrist, J. D. (1998) Estimating Labor Market Impact of Voluntary Military Service Using Social Security Data. *Econometrica*, 66, 249-288.

Angrist, J. D. and A. B. Krueger (1999) Empirical Strategies in Labor Economics. In O. Ashenfelter und D. Card (eds.), *Handbook of Labor Economics*, Vol. III A, chapter 23, 1277-1366.

Börsch-Supan, A. and Hajivassiliou, V.A. (1993) Smooth Unbiased Multivariate Probabilities Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models. *Journal of Econometrics*, 58, 347-368.

Brodaty, T., B. Crepon, and D. Fougere (2001) Using matching estimators to evaluate alternative youth employment programmes: Evidence from France, 1986-1988. In *Econometric Evaluation of Labour Market Policies* (eds. M. Lechner and F. Pfeiffer), Heidelberg: Physica, 85-123.

Dehejia, R. H. and S. Wahba (1999) Causal Effects in Non-experimental Studies: Reevaluating the Evaluation of Training Programmes. *Journal of the American Statistical Association*, 94, 1053-1062.

Dorsett, R. (2001) The New Deal for Young People: relative effectiveness of the options. *Mimeo, Policy Studies Institute*.

Fay, Robert G. (1996) Enhancing the Effectiveness of Active Labour Market Policies: Evidence from Programme Evaluations in OECD countries. *Labour Market and Social Policy Occasional Papers*, #18, OECD.

Frölich, M., A. Heshmati, and M. Lechner (2000) A Microeconometric Evaluation of Rehabilitation of Long-term Sickness in Sweden. *Discussion paper*, 2000-04, *University of St. Gallen*.

Gerfin, M. and M. Lechner (2000) Microeconometric Evaluation of the Active Labour Market Policy in Switzerland. *Discussion paper 2000-10, University of St. Gallen*.

Geweke, J., M. Keane, and D. Runkle (1994) Alternative Computational Approaches to Inference in the Multinomial Probit Model. *Review of Economics and Statistics*, 1994, 609-632.

Heckman, J. J. , H. Ichimura, and P. E. Todd (1997) Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *The Review of Economic Studies*, 64, 605-654.

Heckman, J. J., H. Ichimura, and P. Todd (1998) Matching as an Econometric Evaluation Estimator. *The Review of Economic Studies*, 65, 261-294.

Heckman, J. J., H. Ichimura, J. A. Smith, and P. Todd (1998) Characterisation Selection Bias Using Experimental Data. *Econometrica*, 66, 1017-1098.

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999) The Economics and Econometrics of Active Labour Market Programs. In O. Ashenfelter and D. Card (eds.): *Handbook of Labor Economics*, Vol. III A, chapter 31, 1865-2097.

Heckman, J. J., and R. Robb (1986) Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes. In H. Wainer (ed.), *Drawing Inferences from Self-Selected Samples*, New York: Springer, 1986, 63-107.

Heckman, J. J., and J. A. Smith (1999) The Pre-programme earnings dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies. *The Economic Journal*, 109, 313-348.

Holland, P. W. (1986) Statistics and Causal Inference. *Journal of the American Statistical Association*, 81, 945-970, with discussion.

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

Lalive, R., J. C. van Ours, and J. Zweimüller (2000) The Impact of Active Labor Market Policies and Benefit Entitlement Rules on the Duration of Unemployment. *Mimeo*.

LaLonde, R. J. (1986) Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *American Economic Review*, 76, 604-620.

Larsson, L. (2000) Evaluation of Swedish youth labour market programmes. *Discussion paper, 2000:1, Office for Labour Market Policy Evaluation, Uppsala*.

Lechner, M. (1999) Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany After Unification. *Journal of Business & Economic Statistics,* 17, 74-90.

Lechner, M. (2000a) An Evaluation of Public Sector Sponsored Continuous Vocational Training Programs in East Germany. *The Journal of Human Resources*, 35, 347-375.

Lechner, M. (2000b) Programme Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labour Market Policies. Forthcoming in the *Review of Economics and Statistics*.

Lechner, M. (2001a) Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In M. Lechner and F. Pfeiffer (eds., 2001), *Econometric Evaluation of Labour Market Policies*, Heidelberg: Physica, 43-58.

Lechner, M. (2001b) A note on the common support problem in applied evaluation studies. *Discussion paper 2001-01*, *Department of Economics, University of St. Gallen*.

Lechner, M., and F. Pfeiffer (eds., 2001) *Econometric Evaluation of Labour Market Policies*, Heidelberg: Physica.

Rosenbaum, P. R., and D. B. Rubin (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41-50.

Roy, A. D. (1951) Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3, 135-146.

Rubin, D. B. (1974) Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66, 688-701.

Smith, J., and P. E. Todd (2000) Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators, Forthcoming in the *Journal of Econometrics*.

Spanos, Aris (1999) *Probability Theory and Statistical Inference*. Cambridge: Cambridge University Press.

Tables

Table 1: A matching protocol for the estimation of $\theta_0^{m,l}$

Step 1	Specify and estimate a multinomial probit model to obtain $[\hat{P}_N^0(x), \hat{P}_N^1(x), ..., \hat{P}_N^M(x)]$.
Step 2	Restrict sample to common support: Delete all observations with probabilities larger than the smallest maximum and smaller than the largest minimum of all subsamples defined by *S*.
Step 3	Estimate the respective (counterfactual) expectations of the outcome variables. For a given value of *m* and *l* the following steps are performed: a) Choose one observation in the subsample defined by participation in *m* and delete it from that pool. b) Find an observation in the subsample of participants in *l* that is as close as possible to the one chosen in step a) in terms of $[\hat{P}_N^m(x), \hat{P}_N^l(x), \tilde{x}]$. $\tilde{x}$ contains information on *sex, duration of unemployment*, *native language*, and *start of programme*. 'Closeness' is based on the Mahalanobis distance. Do not remove that observation, so that it can be used again. c) Repeat a) and b) until no participant in *m* is left. d) Using the matched comparison group formed in c), compute the respective conditional expectation by the sample mean. Note that the same observations may appear more than once in that group.
Step 4	Repeat Step 3 for all combinations of *m* and *l*.
Step 5	Compute the estimate of the treatment effects using the results of Step 4.

Table 2: Number of observations and selected characteristics of different groups

Group		obs.	duration of programme	unemployment before...	qualification	foreign	employed March 1999
		(persons)	(mean days)		(mean)	(share in %)	
NONPARTICIPATION	(NONP)	6735	0	250*	1.8	47	39
COMPUTER COURSES	(COC)	1394	36	214	1.3	22	44
EMPLOYMENT PROGRAMMES	(EP)	2473	147	300	1.8	46	26
TEMPORARY WAGE SUBSIDY	(TEMP)	4390	114	228	1.7	46	48

Note: Qualification is measured as skilled (1), semiskilled (2), and unskilled (3). *) Start date simulated.

Table 3: Descriptive statistics of the predicted probabilities from the MNP

	Mean in %	std * 100	correlations			
			non-participation	computer course	employment programme	temporary wage subsidy
Nonparticipation	44.9	12.98	1	-.21	-.48	-.52
Computer course	9.3	8.55		1	-.32	-.19
Employment programme	16.4	11.47			1	-.22
Temporary wage subsidy	29.3	10.88				1

Table 4: Loss of observations due to common support requirement

	Nonparticipation	computer course	employment programme	temp. wage subsidy
Observations before	6735	1394	2473	4390
Observations after	6575	1375	2419	4258
Percent deleted	3	1	2	3

Note: The total number of observations decreases by 365 due to the enforcement of the common support requirement.

Table 5: Share of the largest 10% of the weights to total weight (number of participants) in %

	Nonparticipation	computer course	employment programme	temp. wage subsidy
Nonparticipation		41	35	27
Computer course	21		33	24
Employment programme	24	42		24
Temporary wage subsidy	24	42	35	

Note: Observations from the sample denoted in the column are matched to observations of the sample denoted in the row.

Table 6: Average effects for participants ($\theta_0^{m,l}$) measured as difference in employment rates one year after start of the programme in %-points

l / m	Nonparticipation	computer course	employment programme	temp. wage subsidy
Nonparticipation	40.7	2.1 (3.2)	**7.2** (2.3)	**-6.4** (1.6)
Computer course	**-8.3** (2.5)	45.9	-2.1 (3.5)	**-9.1** (2.7)
Employment programme	**-8.4** (2.3)	-6.5 (4.1)	30.9	**-15.7** (2.5)
Temporary wage subsidy	*4.2* (1.7)	**8.0** (3.3)	**13.8** (2.7)	50.1

Note: Standard errors in parentheses. Results are based on matched samples. **Bold** numbers indicate significance at the 1% level (2-sided test), numbers in *italics* indicate significance at the 5% level. Unadjusted levels on the main diagonal.

Table 7: Average effects of NONPARTICIPATION *for nonparticipants ($\theta_0^{NP,l}$) one year after start: Start dates for nonparticipants*

	Computer course	employment programme	temporary wage subsidy
Baseline	2.1 (3.2)	**7.2** (2.3)	**-6.4** (1.6)
Predicted with covariates	2.5 (2.9)	**8.5** (2.5)	**-4.2** (1.5)
Predicted with covariates and reduced sample	2.9 (3.2)	**8.8** (3.0)	**-5.2** (1.7)

Note: See note on Table 6.

Table 8: Average effects of COMPUTER COURSES *for participants in* COMPUTER COURSES *($\theta_0^{COC,l}$) one year after start: Reduction of information*

	Nonparticipation	employment programme	temporary wage subsidy
Baseline	**-8.3** (2.5)	-2.1 (3.5)	**-9.1** (2.7)
and no long term employment history	**-7.8** (2.5)	1.0 (3.4)	**-8.8** (2.7)
and no duration of current unemployment spell	**-8.9** (2.5)	4.8 (3.3)	**-7.0** (2.7)
and no subjective information	*-5.0* (2.5)	*7.1* (3.3)	**-9.3** (2.7)
and no information on current unemployment spell	-4.1 (2.5)	*7.9* (3.2)	**-8.8** (2.7)
and no information on previous employment, occupation, and skill	1.4 (2.5)	**14.1** (3.1)	**-10.5** (2.6)
and no regional information	-4.6 (2.5)	**14.1** (3.0)	-5.1 (2.7)
Only age, gender, and marital status (no nationality)	3.9 (2.4)	**14.7** (2.8)	**-9.7** (2.2)
No covariates (unadjusted differences)	**5.2** (1.7)	**15.0** (2.1)	*-4.2* (1.9)

Note: See note below Table 6.

Table 9: Average effects of COMPUTER COURSES *for participants in* COMPUTER COURSES ($\theta_0^{COC,1}$)

one year after start: First step

	Nonparticipation	employment programme	temporary wage subsidy
Baseline (200 draws, all 3 correlations between programmes)	-8.3 (2.5)	-2.1 (3.5)	-9.1 (2.7)
2 draws	-8.5 (2.5)	0.9 (3.5)	-8.3 (2.7)
800 draws	-9.6 (2.5)	-0.8 (3.6)	-9.2 (2.7)
3 correlation between NONP and (TEMP, EP, COMP)	-9.1 (2.5)	-1.9 (3.5)	-9.1 (2.7)
Only correlation between TEMP and EP	-5.3 (2.6)	-0.9 (3.5)	-10.3 (2.7)
Only correlation between COMP and EP	-6.9 (2.5)	-1.9 (3.5)	-11.5 (2.7)
Only correlation between COMP and TEMP	-3.3 (2.6)	-2.3 (3.5)	-9.7 (2.7)

Note: See note on Table 6.

Table 9b: Average effects of TEMPORARY WAGE SUBSIDY *for participants in* TEMPORARY WAGE SUBSIDY ($\theta_0^{TEMP,1}$) *one year after start: First step*

	Nonparticipation	computer courses	employment programme
Baseline	4.2 (1.7)	8.0 (3.3)	13.8 (2.7)
No common support	2.3 (1.7)	7.5 (3.3)	13.5 (2.7)
Stricter common support requirement	3.9 (1.8)	8.1 (3.3)	14.9 (2.7)

Note: See note on Table 6.

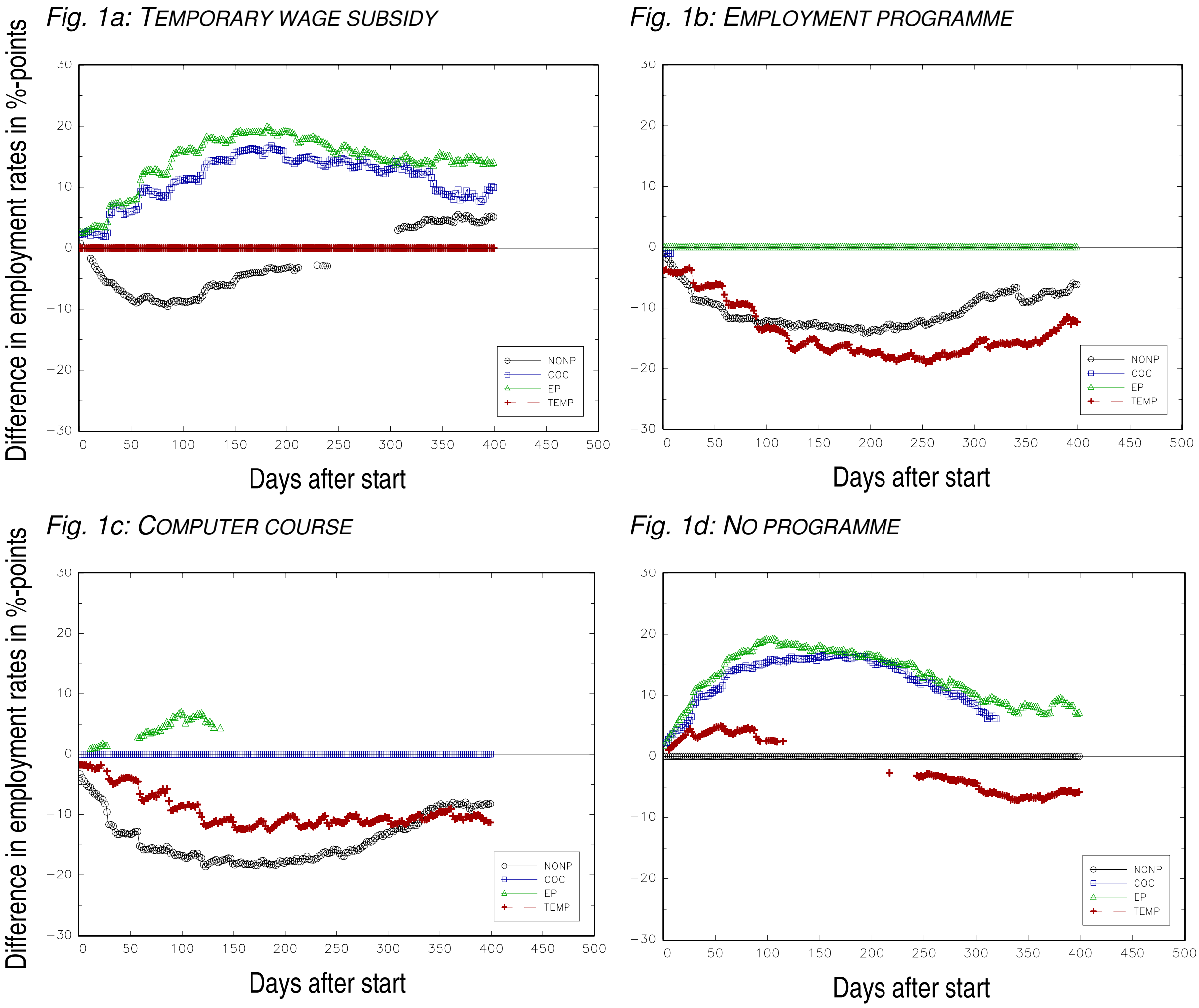
Table 10: Average effects of TEMPORARY WAGE SUBSIDY *for participants ($\theta_0^{TEMP,l}$) one year after start: Bootstrap*

	Approximation									
			Quantiles in %							
	$\hat{\theta}_N^{TEMP,l}$	Std.	2.5	5	25	Median	75	95	97.5	
Nonparticipation	4.2	1.7	0.9	1.4	3.0	4.2	5.4	7.0	7.5	
Computer course	8.0	3.3	1.5	2.6	5.8	8.0	10.2	13.4	14.5	
Employment progr.	13.8	2.7	8.5	9.4	12.0	13.8	15.6	18.2	19.1	
	Bootstrap									
	Mean	Std	Quantiles in %							Normality
	$\hat{\theta}_{N,h}^{TEMP,l}$	$\hat{\theta}_{N,h}^{TEMP,l}$	2.5	5	25	Median	75	95	97.5	p-val. x 100
Nonparticipation	4.3	2.0	0.0	1.2	3.1	4.2	5.8	7.6	8.4	4
Computer course	8.0	3.5	1.1	1.7	5.6	8.1	10.4	13.5	14.2	32
Employment progr.	13.8	2.9	8.3	8.9	11.9	14.0	15.8	19.0	19.2	17

Note: See note on Table 6. 400 bootstrap samples. The bootstrap quantiles are based on the empirical order statistic ($\hat{\theta}_{N,h}^{TEMP,l}$). Normality is tested by the skewness-kurtosis statistic that is asymptotically distributed as $\chi^2(2)$ and attributed to Fisher (see for example Spanos, 1999, p. 745).

Figures

Figure 1: Dynamics of average effects for participants after the start of the programme



Note: NONP: Nonparticipation, COC: Computer course; EP Employment programme; TEMP: Temporary wage subsidy. Only estimated effects that are significant at the 5% level are reported.

Figure 2: Dynamics of average effects of NONPARTICIPATION *for nonparticipants* ($\theta_0^{NP,l}$)

Fig 1d: Baseline

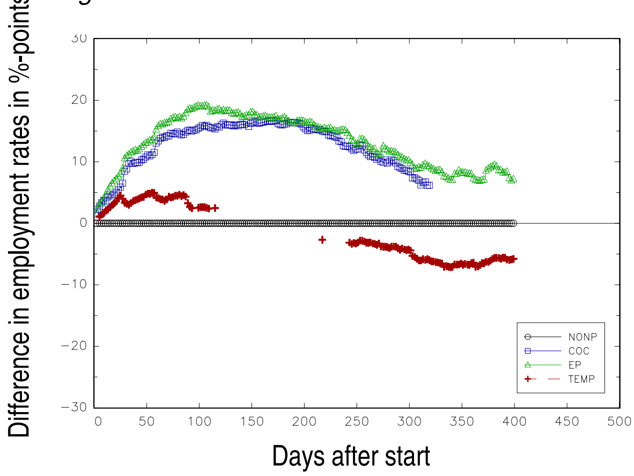


Fig 2a: Predicted with covariates

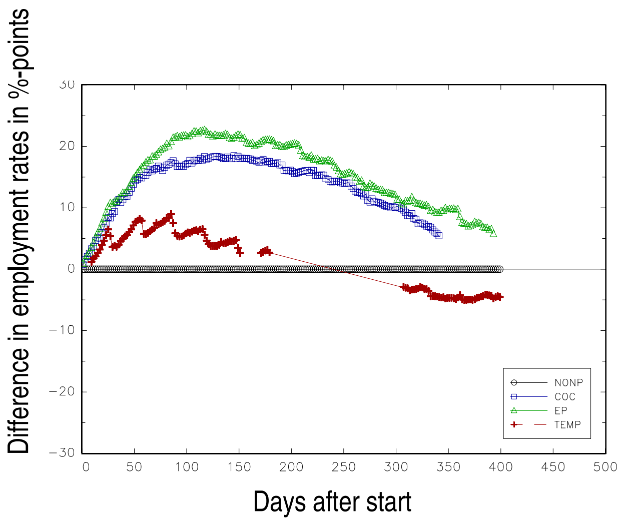
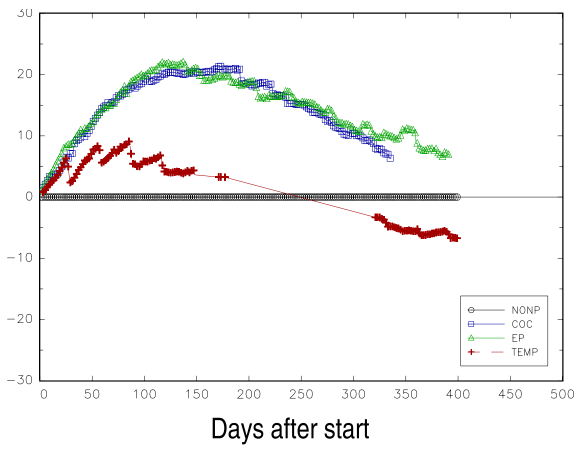


Fig 2b: Predicted with covariates in reduced sample



Note: NONP: Nonparticipation, COC: Computer course; EP Employment programme; TEMP: Temporary wage subsidy. Only estimated effects that are significant at the 5% level are reported.

Figure 3: Dynamics of average effects of COMPUTER COURSES *for participants in* COMPUTER COURSES *(* $\theta_0^{COC,l}$ *): Reduction of information*

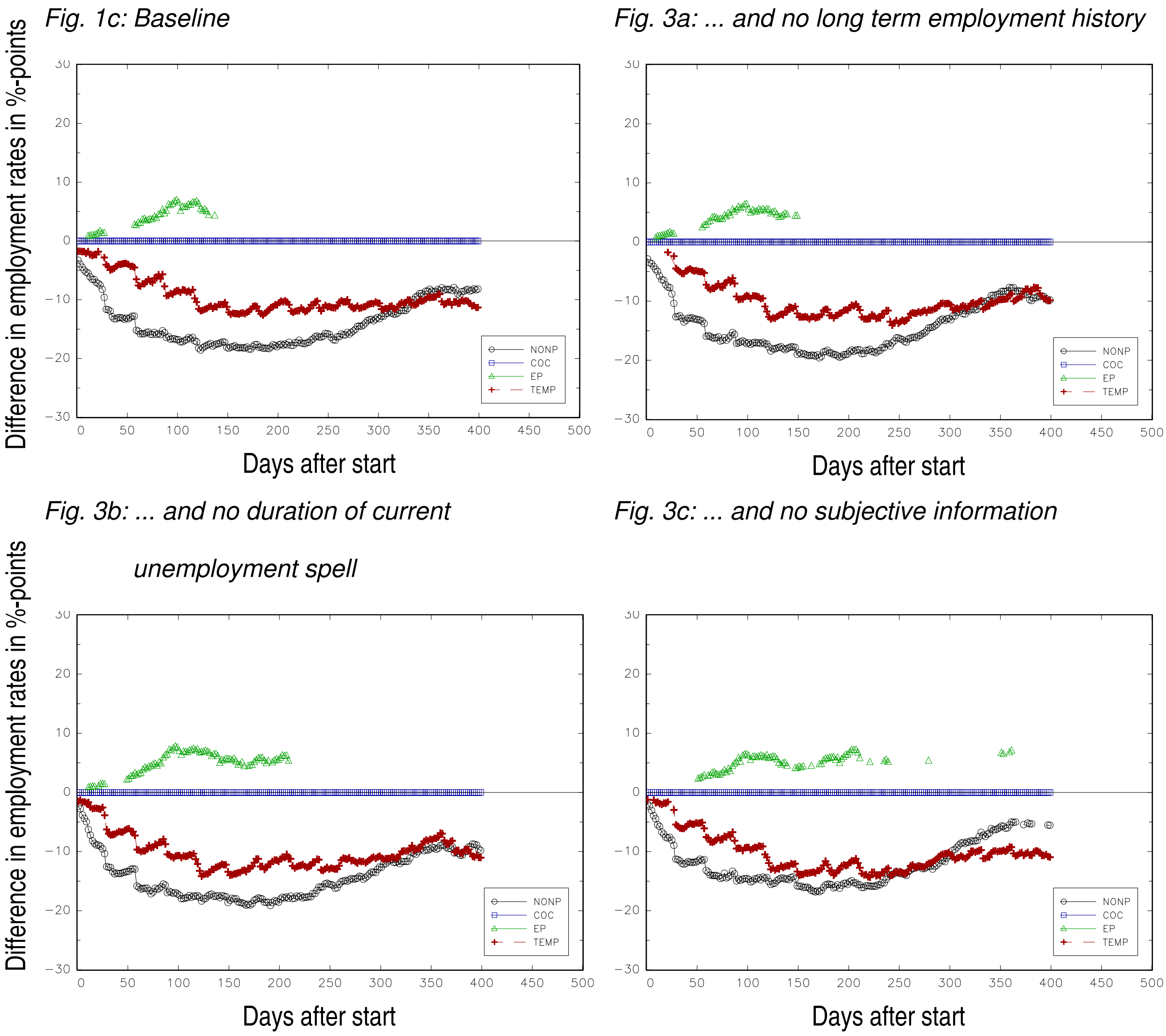


Figure 3 to be continued.

Figure 3 continued

Fig. 3d: and no information on current unemployment spell

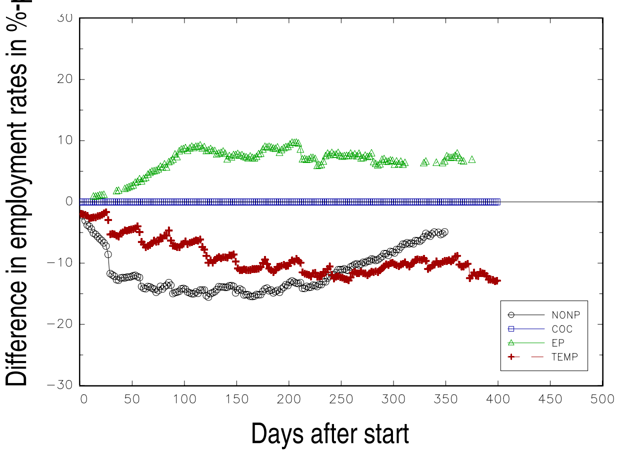


Fig. 3e: ... and no information on previous employment, ...

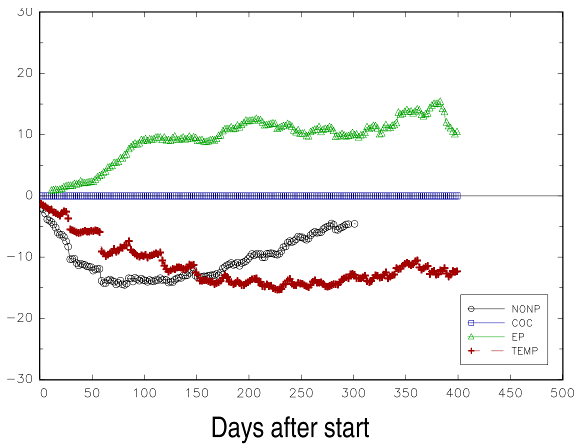


Fig. 3f: ... and no regional information

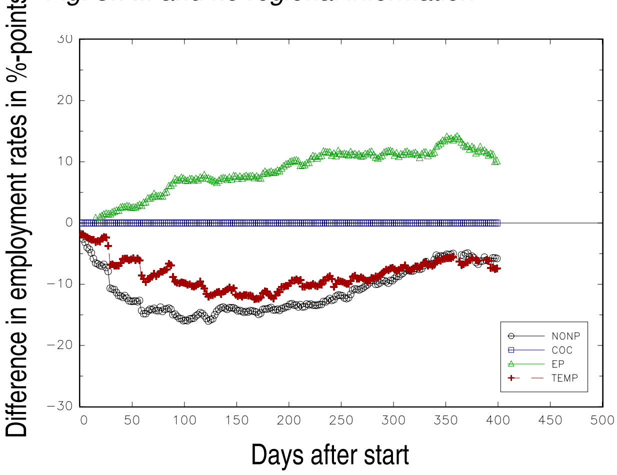


Fig. 3g: Only age, gender, and marital status

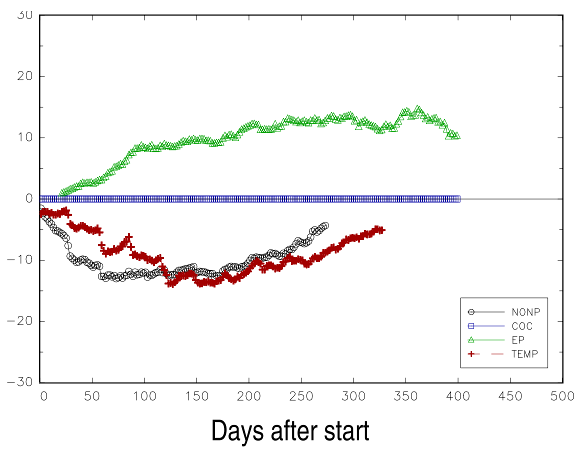
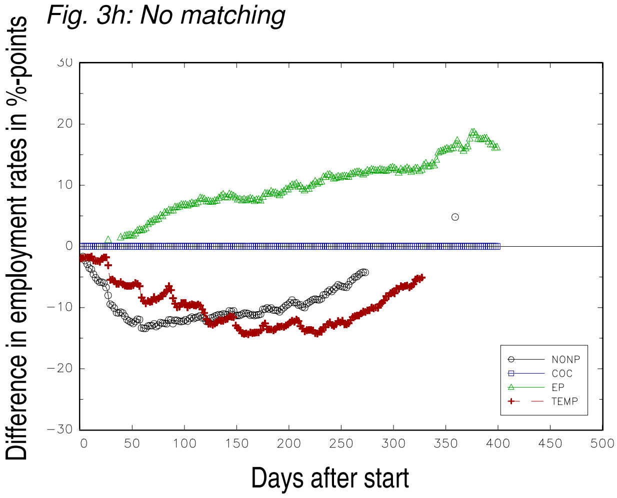


Fig. 3h: No matching



Note: NONP: Nonparticipation, COC: Computer course; EP Employment programme; TEMP: Temporary wage subsidy. Only estimated effects that are significant at the 5% level are reported.

Figure 4: Dynamics of average effects of COMPUTER COURSES *for participants in* COMPUTER COURSES *(* $\theta_0^{COC,1}$ *): MNP estimation in first step*

Fig. 1c: Baseline (all correlations, 200 draws)

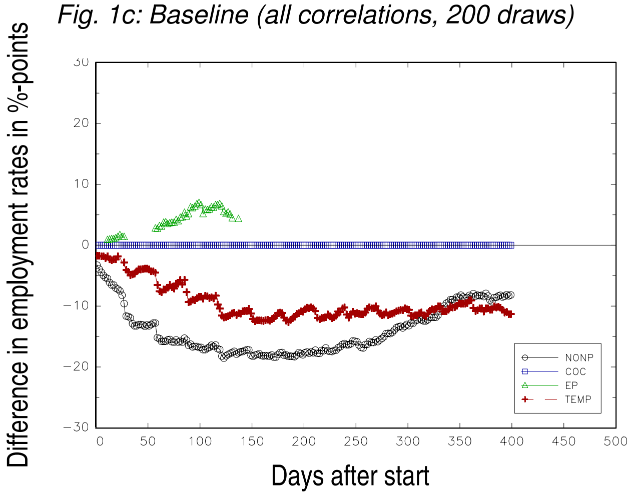


Fig. 4a: 2 draws in GHK simulator

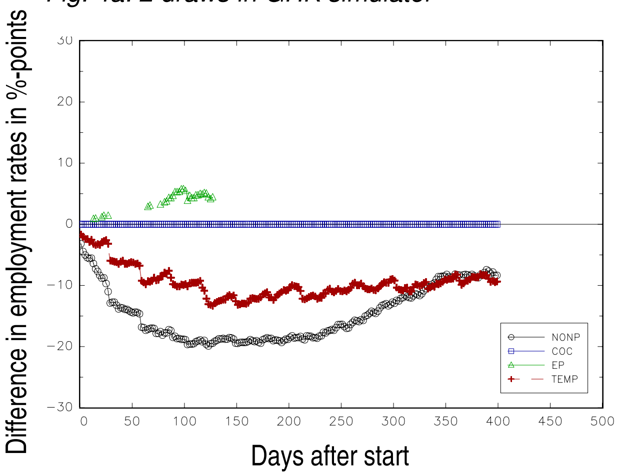


Fig. 4b: 800 draws in GHK simulator

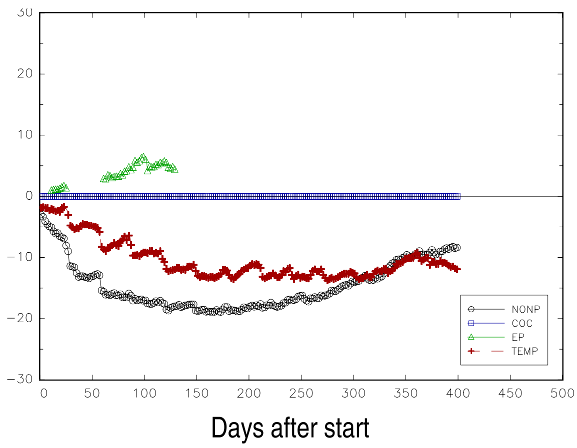


Figure 4 to be continued.

Figure 4 continued

Fig. 4c: Correlation between COMP and EP

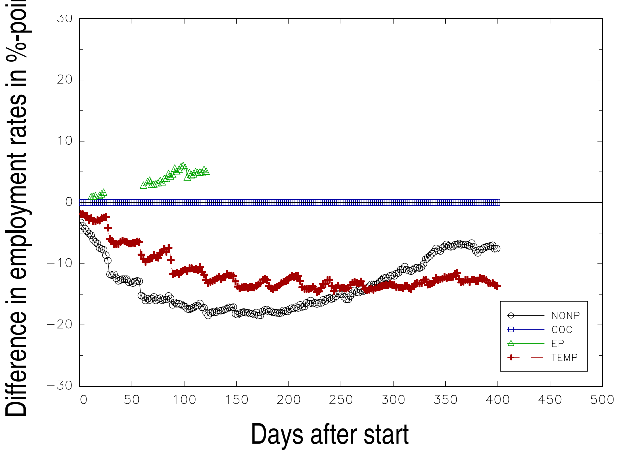


Fig. 4d: Correlation between TEMP and EP

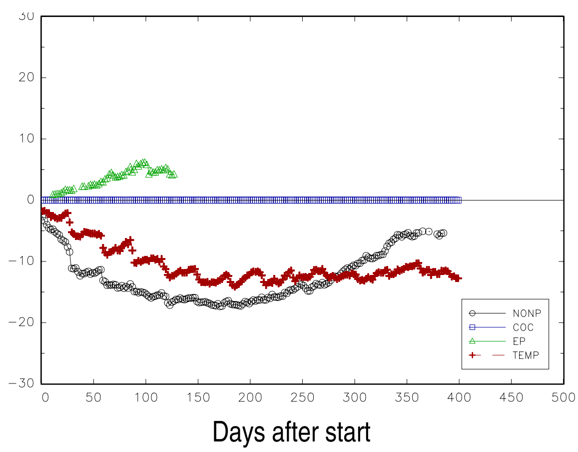


Fig. 4e: Correlation between COMP and TEMP

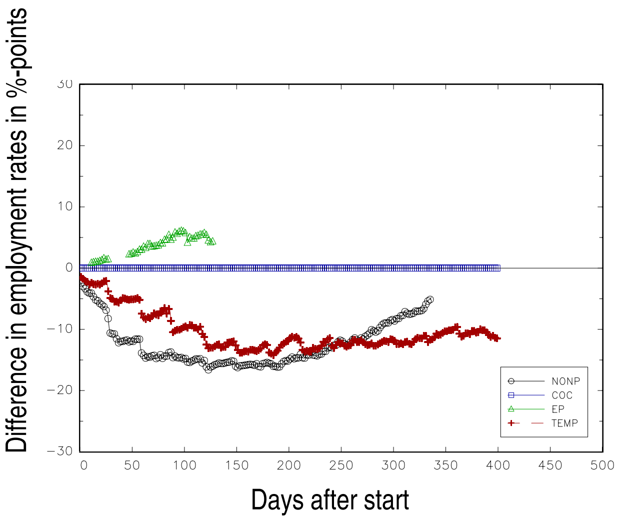
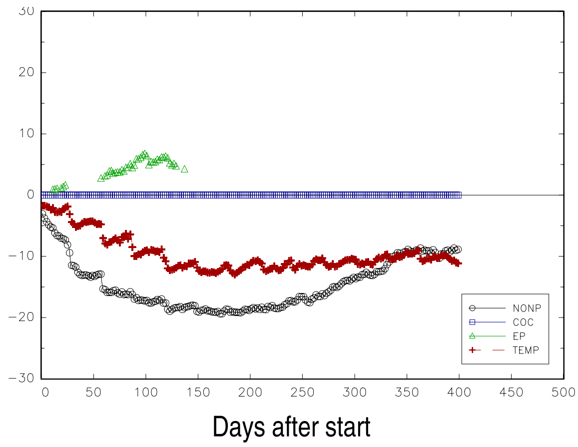


Fig. 4f: Mutual correlations between NONP and COMP, TEMP, EP



Note: NONP: Nonparticipation, COC: Computer course; EP Employment programme; TEMP: Temporary wage subsidy. Only estimated effects that are significant at the 5% level are reported.

Figure 5: Dynamics of average effects of TEMPORARY WAGE SUBSIDY *for participants in*

TEMPORARY WAGE SUBSIDY ($\theta_0^{TEMP,l}$): *Common support*

Fig 1a: Baseline

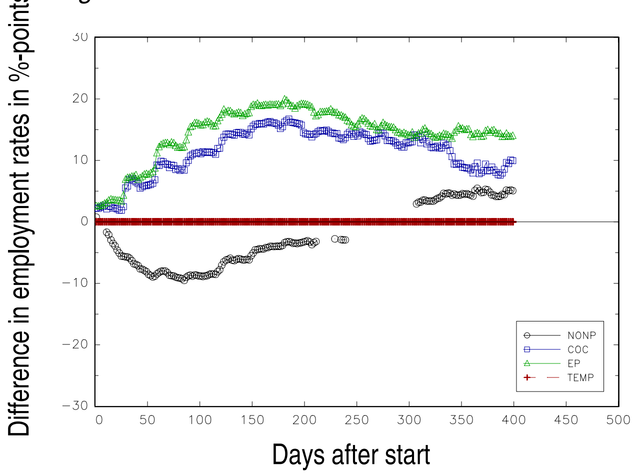


Fig. 5a: No common support required

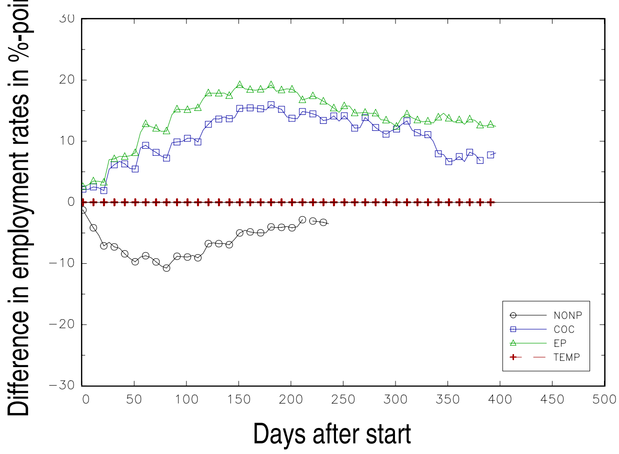
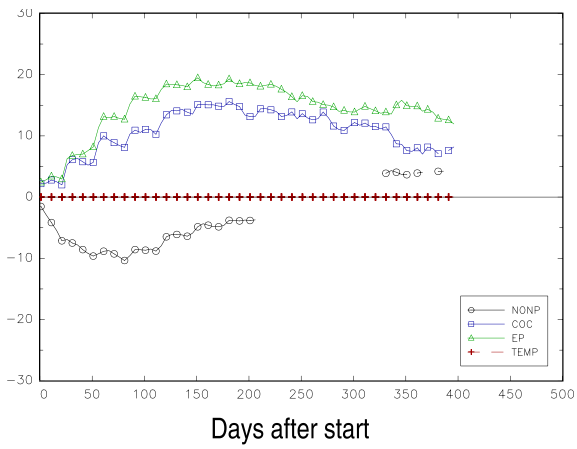
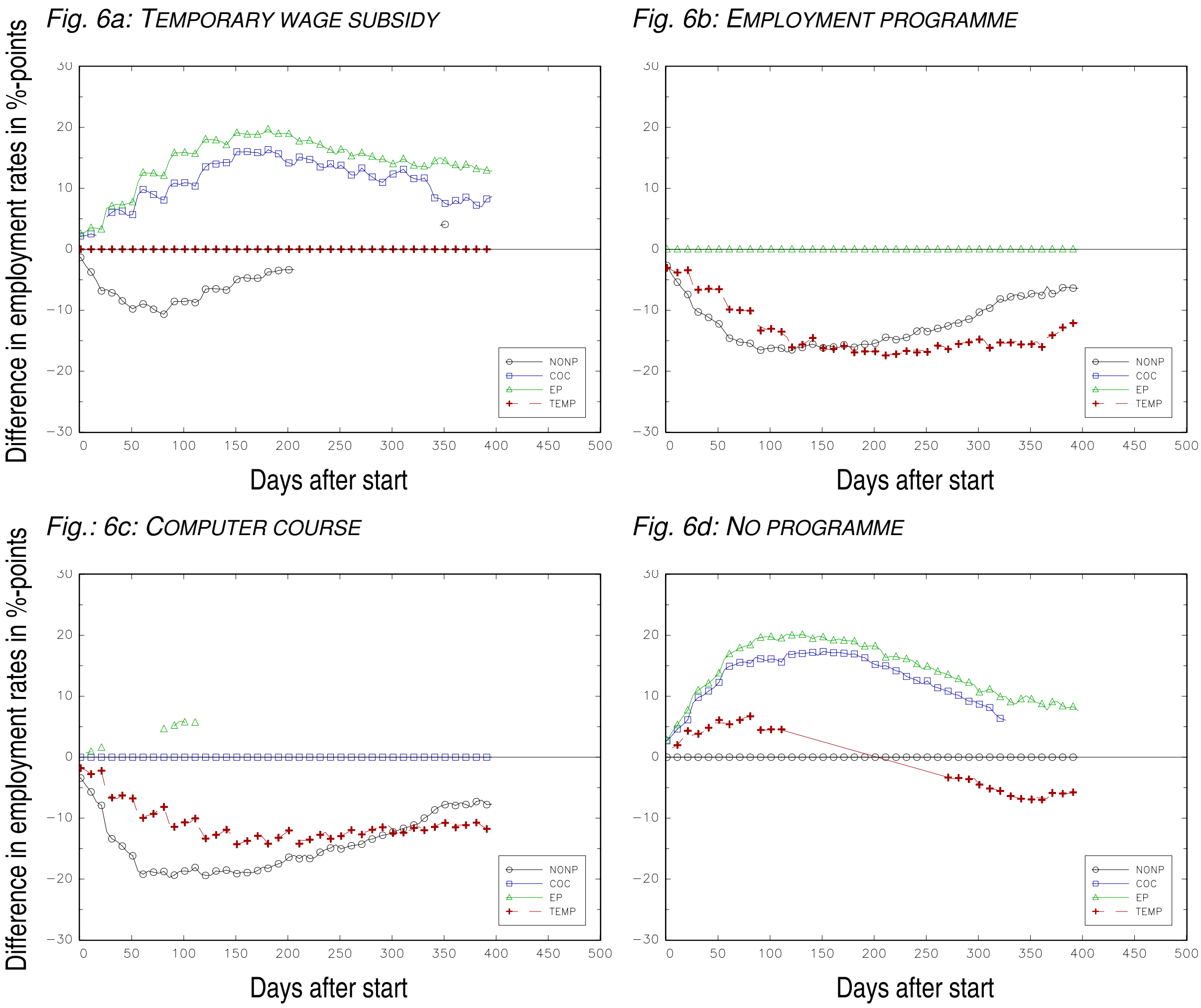


Fig. 5b: Stricter common support required



Note: NONP: Nonparticipation, COC: Computer course; EP Employment programme; TEMP: Temporary wage subsidy. Only estimated effects that are significant at the 5% level are reported.

Figure 6: Dynamics of average effects for participants after the start of the programme: Bootstrap



Note: NONP: Nonparticipation, COC: Computer course; EP Employment programme; TEMP: Temporary wage subsidy. Only estimated effects that are significant at the 5% level are reported. Only every fifth day is displayed. Effects are only displayed if the bootstrap bounds of the 95% interval have the same sign. Based on 400 bootstrap samples.

Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods

Michael Lechner
University of St. Gallen
Swiss Institute for International Economics and Applied Economic Research (SIAW)

Additional appendices containing more information on the data and the estimation of the selection equation

(can be downloaded from **www.siaw.unisg.ch/lechner/l_jrss_a**)

Appendix WWW

Appendix A: Descriptive statistics of the variables used

TableA.1: Descriptive Statistics

Variable	Nonpart.	computer courses	employment programmes	temp. wage subsidy
Number of observations	6735	1394	2473	4390
	Days, Years, Swiss Francs			
Current Unemployment Spell				
Begin of first programme [a)]	96[b)]	80	135	100
Duration of first programme	-	36	147	114
Duration of current unemployment spell at begin of programme	250	214	300	228
Remaining time of benefit entitlement at start of programme	329	410	338	343
Duration of current unempl., 31.12.97	155	134	165	128
Remain. days of „passive regime", 31.12.97	46	59	31	52
Unemployment benefit per day	124.3	124.3	124.9	123.9
Age in years	38.1	38.3	38.8	37.5
	Proportions in %			
Younger than 30	23	23	22	25
Older than 50	11	13	12	9
Female	43	46	48	42
Number of persons to support	2.22	2.22	2.23	2.23
At least one person to support	62	63	63	62
Mother tongue				
German	30	49	33	35
French	21	28	18	19
Italian	12	8	11	12
Not German/French/Italian	37	15	38	34
Language spoken in canton of residence	52	76	52	54
G/F/I, but not canton language	11	10	10	12
Foreign Languages				
Other Swiss language	64	54	67	65
English, Spanish, Portugese	14	26	10	12
Other languages	2	2	1	1
Marital Status				
Single	25	35	29	27
Married	61	47	57	59
Widowed	1	1	1	1
Divorced	13	17	14	13

Table A.1 to be continued

Table A.1 continued

Variable	Nonpart.	computer courses	employment programmes	temp. wage subsidy
Nationality				
Swiss	53	78	54	54
Foreign with permanent permit	32	17	29	31
Foreign with yearly permit	15	5	17	15
Qualification				
Skilled	53	80	48	54
Semi-skilled	16	9	17	17
Unskilled	31	11	35	29
Chances to find a job				
No Information	6	5	5	8
Very easy	6	5	4	6
Easy	13	19	21	17
Medium	53	58	54	56
Difficult	18	11	20	12
Special case	4	1	4	2
Mobility				
Not mobile	13	8	5	8
Daily commuter	83	85	90	88
Mobile within Switzerland or abroad	5	7	5	5
Looking for ..... job				
Full-time	34	34	36	38
Part-time	16	18	12	12
No information	49	49	52	50
Unemployment-status				
Full-time	77	77	82	81
Part-time	19	18	14	13
In part-time employment	2	1	1	4
Other	3	3	2	3
Monthly earnings in last job				
Less than 1000	2	4	2	3
Between 1000 and 2000	11	12	11	11
Between 2000 and 3000	25	22	24	24
Between 3000 and 4000	27	27	27	28
Between 4000 and 5000	20	17	19	20
Between 5000 and 6000	8	9	8	8
More than 6000	7	9	8	7
Duration of unemployment spell at beginning of programme				
Less than 90 days	16	18	7	19
Less than 180Tage	39	48	24	44
Less than 270 days	57	70	43	63
Less than 365 days	76	85	66	81
More than 365 days	24	15	34	19
Job position				
Self-employed	1	2	1	0
High (management, etc.)	6	9	3	5
Medium	56	73	49	58
Low	37	16	47	37

Table A.1 to be continued

Table A.1 continued

Variable	Nonpart.	computer courses	employment programmes	temp. wage subsidy
Previous occupation				
Agriculture	2	1	2	2
Mining	0	0	0	0
Food, tobacco	1	1	1	1
Textiles	1	1	1	1
Wood and paper	1	1	2	1
Chemical	0	0	0	0
Metals	7	6	8	8
Watches, jewelry	0	0	1	0
Health care	3	3	3	3
Architecture, engineers	1	5	1	2
Construction	8	3	8	10
Transportation	4	2	4	4
Restaurants	17	8	14	17
Printing	1	1	1	1
Minerals	0	0	0	0
Entrepreneurs, senior officials, justice	4	5	2	2
Painting, technical drawing	5	6	6	8
Office and computer	14	28	12	12
Retail trade	9	13	7	7
Security, cleaning, clerical, social work	5	2	5	5
Science	2	3	1	1
Artist	2	2	1	2
Education	2	2	2	3
News and communication	1	2	1	1
Body care	1	0	0	1
Other	8	5	14	8
Correspondence between desired and previous job				
2-digit	73	74	69	75
3-digit	68	66	63	69
Previous industry sector				
Agriculture	1	1	2	2
Mining, energy, water	0	0	0	0
Construction	13	7	12	17
Public services	11	9	10	6
Other services	5	6	6	5
Health care	4	4	3	4
Research and development	0	1	0	0
Education	2	2	2	2
Banking, insurance	3	6	2	2
Real estate	1	1	1	1
Consulting	11	16	11	12
Transportation	3	4	3	3
News and communication	3	1	0	0
Trade	14	19	15	13

Table A.1 to be continued

Table A.1 continued

Variable	Nonpart.	computer courses	employment programmes	temp. wage subsidy
Restaurants, catering	15	8	12	16
Repairs	2	2	1	2
Food, tobacco	1	1	1	1
Textiles	1	1	1	1
Wood, furniture	1	1	1	1
Paper, paper products	0	0	0	0
Printing	1	3	1	1
Leather	0	0	0	0
Chemical	1	1	1	1
Non-ferrous minerals	1	0	1	1
Metals	2	2	4	2
Machinery and equipment	2	2	2	3
Electrical machinery, optics	2	2	2	2
Watches, jewelry	1	1	1	1
Other manufacturing	1	1	1	1
Industry unemployment rate in %, 1/98	6.6	5.7	6.3	6.7
Canton				
Zurich	21	21	17	18
Berne	7	9	14	10
Lucerne	2	5	3	3
Uri	0	0	0	0
Schwyz	0	1	1	1
Obwalden	0	0	0	0
Nidwalden	0	0	0	0
Glarus	0	0	0	0
Zug	1	1	1	1
Freiburg	3	4	4	3
Solothurn	2	2	5	4
Basel-City	3	4	3	3
Basel-Landschaft	2	3	2	2
Schaffhausen	1	2	1	1
Appenzell AR	0	0	1	0
Appenzell IR	0	0	0	0
St. Gall	4	4	2	5
Graubünden	1	2	2	1
Aargau	5	4	5	5
Thurgau	1	3	2	2
Ticino	9	5	8	8
Waadt	15	17	12	13
Wallis	4	3	7	8
Neuenburg	4	2	5	3
Geneva	12	9	4	7
Jura	1	1	1	1
Cantonal unemployment rate	5.35	5.09	5.09	5.21

Table A.1 to be continued

Table A.1 continued

Variable	Nonpart.	computer courses	employment programmes	temp. wage subsidy
Canton language				
German	53	59	59	58
French	38	36	33	34
Italian	9	5	8	8
Region				
Eastern	8	10	7	10
Central	4	7	7	5
South-west	31	29	23	28
North-west	10	10	10	11
West	17	18	29	20
Size of town where worked before				
<1000	9	9	8	9
<2000	16	16	17	18
<5000	32	31	35	36
<10'000	44	45	49	49
<20'000	62	62	65	67
<30'000	67	68	71	73
<50'000	72	73	77	77
<100'000	76	78	81	81
> 100'000.	24	22	19	19
<200'000 .	92	93	93	94
> 200'000	8	7	7	6
Region of placement office				
Large city	47	45	35	38
Small city	37	40	46	42
Rural	16	14	19	19
No information	1	1	1	1
Inflow to long-term unemployment [c)]	2.1	2.1	2.0	1.9
Outflow from long-term unemployment [d)]	1.1	1.2	1.1	1.1
No information	17	18	20	21
Remaining benefit eligibility				
Less than 6 months	22	14	16	20
Less than 12 months	49	32	50	45
Less than 18 months	78	62	82	72
More than 18 months	15	28	13	19
Unemployment history				
First spell	60	67	63	57
Number of spells prior to current spell	0.51	0.39	0.46	0.55
Duration of previous spell / 1000	0.08	0.07	0.09	0.08

Table A.1 to be continued

Table A.1 continued

Variable	Nonpart.	computer courses	employment programmes	temp. wage subsidy
Sanction days without benefit payment				
Number of sanction days during last unemployment spell	4.2	3.6	4.7	3.7
Share in total unemployment spell	0.05	0.05	0.06	0.05
Positive number of sanction days (in %)	25	22	24	22
Previous programme participation				
Sum of short programs between July and December 1997	0.04	0.09	0.06	0.06
Participation in training course or employment programme between July and December 97 (less than 14 days)	1	0	1	0
Employment programme before July 97	1	1	1	1
Training course before July 97	1	0	1	0
Temporary wage subsidy before July 97	1	1	1	3
Employment history from social security data				
Number of months unemployed since entry into social security system	7.6	6.4	8.2	6.4
Number of months employed since entry into social security system	85	91	84	90
Number of months out of labour force since entry into social security syst.	15.5	14.0	15.3	12.6
Never unemployed	37	43	37	40
Month of entry into social security system	12.0	8.7	12.7	11.4
Number of employment spells	3.53	3.03	3.53	3.51
Number of unemployment spells	1.41	1.11	1.46	1.29
Mean duration of employment spell in months	40	49	39	43
Mean duration of unemployment spell [e] in months	5.9	6.3	6.2	5.3
Standard deviation of wages / 1000	0.99	1.06	0.90	0.94
Duration of last employment spell	40	48	40	43
Wage growth during last employment spell	81	113	69	69
Proportion of time unemployed in %	7	6	8	6
Proportion of time employed in %	78	81	77	82

Notes: [a] The begin of a programme is measured in days since 1.1.98. [b] Simulated.
[c] Mean number of transition into long-term unemployment relative to total unemployment within regional placement offices. [d] Mean number of transitions to employment relative to total unemployment within regional placement offices
[e] This variable takes a value of zero, if person has never been unemployed before.

Appendix B: Estimates of the multinomial probit model for the baseline scenario

Table B.1: Estimated coefficients of a multinomial probit model for participation in a programme

Variable	computer courses	employment programme	temporary wage subsidy
Age in years / 10		***0.15***	
Female	-0.04	***-0.26***	***0.21***
Marital status married	***-0.13***	***-0.28***	
Marital status divorced			*0.08*
Mother tongue			
Not German/French/Italian	***-0.25***		***-0.27***
G/F/I, but not canton language		***-0.21***	***-0.16***
Foreign Languages			
Other Swiss language	***0.20***	***0.16***	***0.17***
English, Spanish, Portuguese	***0.28***		
Looking for ... job (reference category: no information and full-time)			
Part-time			***-0.23***
Unemployment-status (reference category: part-time)			
Full-time	**0.14**	***0.40***	***0.23***
In part-time employment			***1.16***
Nationality (reference category: Swiss)			
Foreign with permanent permit	***-0.31***	***-0.16***	
Foreign with yearly permit	***-0.41***	-0.04	
Monthly earnings in last job			
Less than 2000	**0.13**		
Chances to find a job (reference category: medium)			
No information	*-0.16*	***-0.36***	0.08
Very easy	0.05	***-0.25***	0.00
Easy	0.08	***-0.15***	**0.13**
Difficult	**-0.16**	0.01	***-0.33***
Special case	***-0.53***	-0.02	***-0.86***
Qualification (reference categories: semi-skilled, unskilled)			
Skilled	***0.36***		
Previous industry sector			
Construction		***-0.25***	
Public services			***-0.34***
Consulting	***0.20***		
Restaurants, catering		***-0.49***	
Printing	***0.44***		
Job position function (reference category: assistant)			
Self-employed			***-0.71***
High (management, etc.)	*0.16*	***-0.62***	
Medium	***0.24***	***-0.20***	

Table B.1 to be continued

Table B.1 continued

Variable	Computer courses	employment programme	temporary wage subsidy
Previous occupation (reference categories: mining, wood and paper, chemical, minerals, artist)			
Agriculture	**-0.54**		
Metals	**-0.16**	**-0.15**	
Architecture, engineer	***0.63***		
Construction	***-0.49***	-0.03	
Transportation	**-0.31**		
Printing			**-0.51**
Entrepreneurs, senior officials, justice		***-0.36***	***-0.75***
Painting, technical drawing			***0.25***
Office and computer	***0.35***		***-0.25***
Retail trade	***0.22***	***-0.24***	***-0.36***
Security, cleaning, clerical, social work	***-0.35***		
Science			***-0.52***
Education	***-0.53***		
News and communication	***0.60***		
Body care	**-0.65**	***-1.21***	
Other		***0.19***	
Desired = previous job, 3-digit	**-0.10**		
Additional regional effects by canton			
Schwyz		***0.99***	
Freiburg		**0.29**	
Solothurn		0.04	
Basel-City		***-0.64***	*-0.30*
St. Gall		***-0.94***	
Graubünden	***0.66***	-0.20	**-0.47**
Aargau	***-0.56***	-0.18	-0.18
Thurgau	**0.36**	0.11	
Ticino	***-1.06***	***0.98***	-0.08
Waadt		***-1.15***	***-0.56***
Neuenburg	***-0.80***		***-0.58***
Geneva	**-0.19**	***-2.10***	***-0.72***
Jura			***-0.66***
Cantonal unemployment rate	***0.20***	***-0.24***	-0.04
Region (reference category: Zurich)			
Eastern	*0.21*	*0.27*	***0.33***
Central	***0.63***	***0.44***	0.06
South-west	***-0.71***	***1.71***	***0.65***
North-west	***0.32***	***0.44***	**0.30**
West	0.07	***0.74***	***0.31***
Size of town where worked before			
>200'000	***-0.33***		
<30'000			**0.11**
<5000		0.02	
<2000	**-0.13**		
Region of placement office			
Large city		-0.08	
Rural	***-0.26***		

Table B.1 to be continued

Table B.1 continued

Variable	Computer courses	employment programme	temporary wage subsidy
Long-term unemployment in regional placement office			
Inflow to long-term unemployment	***2.55***	***2.39***	
Outflow from long-term unemployment	***3.15***	***2.64***	
No information	***1.00***	***0.82***	
Sanction days without benefit payment			
Positive number of sanction days (in %)	**-0.13**		**-0.08**
Unemployment history			
First spell	***0.17***	***0.17***	
Number of spells prior to current spell			
Previous programme participation			
Sum of short programs between July and Dec. 1997	***0.31***	*0.11*	0.09
Employment programme before July 97		*0.27*	
Temporary wage subsidy before July 97			***0.73***
Begin of programme / 100	0.00	***0.34***	***0.26***
Duration of unemployment spell at beginning of programme			
Duration (days)	**-0.80**	**-0.74**	***-2.52***
Less than 90 days	***-0.22***	***-0.45***	*-0.11*
Less than 180 days		***-0.38***	*-0.13*
Less than 270 days		**-0.16**	**-0.15**
Remaining days of "passive regime" on 31.12.97	***0.29***		-0.07
Employment history from social security data			
Never unemployed	***0.23***		
Mean duration of employ. spell in months		**-0.15**	
Mean duration of unemploy. spell in months	***2.90***		
Standard deviation of wages / 1000	***-0.11***	***-0.24***	***-0.13***
Proportion of time unemployed, in %	**-1.02**	***0.96***	**-0.61**
Proportion of time employed, in %			***0.84***

Note: Simulated maximum likelihood estimates using the GHK simulator (200 draws in simulator for each observation and choice equation). Coefficients of the category NONPARTICIPATION are normalised to zero. All equations include a constant. Inference is based on the outer product of the gradient estimate of the covariance matrix of the coefficients ignoring simulation error. *N = 14992*. Value of log-likelihood function: - 16844.4.
Bold and italic numbers indicate significance at the 1% level (2-sided test), **bold** numbers at the 5% level, and numbers in *italics* relate to the 5% level.
If not stated otherwise, all information in the variables relates to the last day in December 1997.

Table B.2: Estimated covariance and correlation matrices of the error terms in the multinomial probit model

	Nonpart.		computer courses		employment programmes		temporary wage subsidy	
	Coef	t-val	coef	t-val	coef	t-val	coef	t-val
				Covariance matrix [a]				
nonparticipation	1		0	-	0	-	0	-
computer courses			1		-0.63	-2.00	-0.37	-1.46
employment programmes					1.40		-0.67	-2.75
temporary wage subsides							1.95	
				Correlation matrix [a] x 100				
nonparticipation	100		0	-	0	-	0	-
computer courses			100		-53.3	-	-26.5	-
employment programmes					100		-40.5	-
temporary wage subsides							100	

Note: [a] Three Cholesky factors are estimated to ensure that the covariance of the errors remains positive definite. t-values refer to the test whether the corresponding Cholesky factor is zero (off-diagonal).